R E P O R T R E S U M E S

ED 003 509 56

TELEVISION IN HEALTH SCIENCES EDUCATION.

BY- GRANT, THEO. S. MERRILL, IRVING R.
CALIFORNIA UNIV., SAN FRANCISCO. MEDICAL CENTER

REPORT NUMBER NDEA-VIIA-064 PUB DATE 30 SEP 63
GRANT OEG-7-14-1440003
EDRS PRICE MF-$0.18 HC-$2.60 69P.

DESCRIPTORS- *HEALTH EDUCATION, *TELEVISED INSTRUCTION, *DENTAL HEALTH, DENTAL CLINICS, *HOSPITAL SCHOOLS, *TELEVISION RESEARCH, TELEVISION TEACHERS, TELEVISION VIEWING, CLOSED CIRCUIT TELEVISION, HEALTH PROGRAMS, VIDEO TAPE RECORDINGS, SAN FRANCISCO, CALIFORNIA

A MAJOR MEDICAL CENTER CONDUCTED A SERIES OF EXPERIMENTAL STUDIES CONCERNED WITH THE USE OF CLOSED CIRCUIT TELEVISION INSTRUCTION IN THE CURRICULUMS OF MEDICINE, DENTISTRY, PHARMACY, AND NURSING. THE SIX STUDIES REPORTED WERE (1) OVER 300 HEALTH SCIENCE TELEVISION PRESENTATIONS WERE PRODUCED, PRESENTED TO STUDENTS, AND EVALUATED. REPORTS WERE MADE AT THE CONCLUSION OF EACH PROGRAM. THE EFFICIENCY OF TELEVISION IN HEALTH SCIENCE EDUCATION IS DESCRIBED IN TERMS OF TELEVISION'S CAPABILITY OF IMAGE MULTIPLICATION, MAGNIFICATION, ASSOCIATION, TRANSFORMATION, AND MEMORY. (2) HOME AND HOSPITAL VIEWING OF CONTINUING EDUCATIONAL BROADCASTS WERE INVESTIGATED UNDER THREE PRESENTATION-RESPONSE CONDITIONS. INDIVIDUAL VIEWING AT HOME WAS FOUND TO BE SUPERIOR TO GROUP VIEWING AT THE HOSPITAL. NO DIFFERENCES WERE FOUND FOR THE PRESENTATION-RESPONSE CONDITIONS, (3) A MEASUREMENT WAS MADE OF ACHIEVEMENT IN SCIENCE REASONING FOR PHYSIOLOGY. THE TELEVISED INSTRUCTOR-DEMONSTRATION APPEARED AS EFFICIENT AS THE LABORATORY EXERCISE, (4) AN INVESTIGATION WAS MADE OF THE DECREASE IN RETENTION OF TELEVISION PRESENTATIONS DURING THE DELAY BETWEEN DENTAL LABORATORY INSTRUCTION AND APPLICATION. NO DIFFERENCES WERE FOUND BETWEEN GROUPS RECEIVING INSTRUCTION BY TELEVISION AND ILLUSTRATED LECTURE, (5) STEP-SIZE WAS INVESTIGATED IN TELEVISION PRESENTATIONS ON DENTAL TECHNIQUES, AND (6) THE EFFICIENCY OF TELEVISION PRESENTATIONS EMPLOYING SUBJECTIVE AND OBSERVER CAMERA ANGLES WAS INVESTIGATED. (TC)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Experimentation and Development of New and More Effective Techniques and Methods for Training Teachers in and Presenting Academic Subject Matter through the Special Media of Television and Video Tapes in the Health Sciences.

Short Title:

TELEVISION IN HEALTH SCIENCES EDUCATION

Report Prepared by:

Theo. S. Grant and Irving R. Merrill

Principal Investigator:
Theo. S. Grant

Title VII Project Number 064
National Defense Education Act of 1958
Grant Number 7-14-1440003

Grantee:
University of California, San Francisco Medical Center

The Research Reported Herein Was Supported by a Grant From the
U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education

September 30, 1963

ED 003509

PREFACE

This is a report of several studies conducted under the only three-year Title VII grant to be concerned with closed-circuit television instruction in a major medical center, including the curricula in medicine, dentistry, pharmacy, and nursing. Because of its scope and exploratory nature, a certain latitude in the development of the directions of the project was considered advisable.

Scope

In general outline, the research was conceived to fall into three phases, each of approximately one year's duration. Phase I was to be devoted to education of the faculty in the use of closed-circuit television. Phase II was to concentrate on the area of undergraduate instruction, and Phase III was to investigate the use of broadcast television in the continuing education of the health science professional. Specific research objectives were indicated within each phase.

As a consequence, the findings represent a number of promising approaches to the use of the medium within the health sciences rather than a comprehensive investigation of a single, restricted problem. The findings include differences of practical as well as statistical significance, particularly where alternative uses of television have been compared.

Phase I was devoted to preparing the instructor to use television to its best advantage and to the development of techniques and equipment suitable to formal research studies (Grant, 1959, p. 6). Several methods of instruction were employed. Practical experience in teaching by television was stressed, as indicated by data in Table 1 for undergraduate professional curricula in all four schools.

Table 1

Undergraduate Instruction by Closed-circuit Television giving Practical Experience in Use of the Medium to Members of the Medical Center Faculty

	Instructors	Television Presentations	Hours of TV Teaching	Student Hours of TV Instruction
1960-61	28	50	89-1/2	7,024-1/2
1961-62	37	63	107	7,599-1/2
1962-63	59	103	137	9,336-1/2
Total:	124	216	333-1/2	23, 960-1/2

A seminar on Television and Undergraduate Instruction on the San Francisco Campus was conducted on October 1, 1960, for department heads or their permanent television representatives. Based upon the practical experience of the faculty in use of the medium during the Fall semester, 1960, a School of Dentistry Workshop in Teaching by Television was conducted during February, 1961. The workshop was limited to six participating faculty members in order to permit maximum experience with equipment, opportunity for constructive group criticism, and personalized instruction.

Prior to the workshop, a Summary of Findings in Instructional Television Research with special attention to studies in the field of the health science professions was distributed to deans, department heads, and faculty representatives for television, as well as workshop participants. An informal survey of the effect of distribution indicated that the summary encouraged several instructors to try closed-circuit television and that it suggested use of alternate television teaching techniques for others.

A major part of the activities of Phase I was the study and analysis of how the health sciences instructor and the television staff could best work together to accomplish

the instructional objectives of the televised presentation. Appendix A summarizes the conclusions on faculty-television staff cooperation drawn from this phase of the project.

Phase II, formal research using the subject matter of the undergraduate professional curricula, included a number of requirements. It was to involve all four schools (Grant, 1959, p. 3); investigate several types of health sciences instructional objectives (Grant, 1959, p. 4); determine "new uses of television in medical education," i.e., the capabilities of the television system in relation to health sciences subject matter, such as image multiplication and image transportation (Grant, 1959, p. 5); and to conduct a study of camera placement, including what constitutes the "optimum advantage point" for the television camera (Grant, 1959, p. 3).

The statement of procedures to be followed in all formal research under the grant excluded investigation of student attitude and prescribed measurement of learning achievement and retention (Grant, 1959, p. 4) and of savings in instructor and/or student time (Grant, 1959, p. 5).

Phase III, formal research using the subject matter of continuing education, was to be devoted to investigation of how television could best allow many more professionals to have the advantage of a program of continuing education without travel from their places of practice (Grant, 1959, p. 3). An investigation of the efficiency of television in reporting health sciences information to the general public was included as a contingency, to be pursued only when, or if, progress permitted (Grant, 1959, p. 4). The use of videotapes and/or kinescopes was prescribed for broadcast presentation in these outlying communities.

Separate Studies

To accomplish the objectives of this project, a number of separate studies were conducted. Each chapter of the monograph concerns a particular study, with the exception of Chapter I which draws upon findings of several of the studies. The chapters may be classified according to curriculum involved, educational level of subject matter, objectives of instruction, and aspects of the television system investigated as follows:

Curriculum involved -- Dental, Chapters IV, V; Health Sciences, Chapter I; Medical, Chapter III; Nursing, Chapters II, VI; Pharmacy, Chapter III.

Educational level -- All levels, Chapter I; Continuing education, Chapters II, VI; Undergraduate education, Chapters III, IV, V.

Instructional objectives -- Performance of perceptual-motor tasks, Chapters IV, V; Recognition learning of specific factual information, Chapter II; Recognition of complex interpersonal behaviors, Chapter VI; Science reasoning, Chapter III.

Television system -- Camera placement, Chapter VI; System capabilities, Chapter I.

It is now fairly clear that the simultaneous presentation of a picture or other image with the spoken word during instruction does not necessarily increase learning achievement; under certain conditions, simultaneous presentation may interfere with learning. Because televised instruction consists almost entirely of picture plus spoken commentary, it is important to be able to specify the conditions under which simultaneous presentation facilitates learning. However, more definitive information on pictorial recognition is needed before studies involving the simultaneous stimulation of two sensory modalities can be most fruitful. Appendix B reports on a study of perceptual recognition of pictorial signs.

Acknowledgements

It is a pleasure to acknowledge the support and encouragement of J. B. deC. M. Saunders, M.B., Ch.B., Provost and Chief Administrative Officer of the University of California Medical Center in San Francisco, because any investigation of a new development in education must have active administrative support if it is to succeed. Not only did he accept the invitation to serve as chairman of several meetings at which the television project was introduced to the faculty, but he did much to facilitate the establishment of new administrative channels necessary to the success of the project. W. C. Fleming, D.D.S., Vice Provost and Dean, School of Dentistry, not only actively participated in person during Phase I of the project, but he also allocated dental school funds as a portion of the University's matching contribution toward the project budget. Because this project differed from most medical research grants, more than routine assistance was required from the business office. Thanks are due S. C. Bateman, Business Manager, for a full measure of support, including assistance in securing the project staff and other administrative matters. His optimism and encouragement have been deeply appreciated.

The inspiration and counsel of D. S. Ruhe, M.D., University of Kansas Medical Center, were an important influence in establishing the scope of the project.

Special thanks are due Associate Investigators L. L. Bennett, M.D., and Betty J. Dietrich, R.N., M.S., who provided wise counsel throughout the project and who actively participated in one or more of the special studies. Without the assistance of J. G. Shrock, Ed.D., and Ruby Y. Buenz the project would not have been completed on schedule. Their loyalty to the research project is greatly appreciated. The sucessful execution of the several studies depended in large measure upon the skill and competence which R. A. Litke and the television staff demonstrated in the preparation of the television presentations. Particular thanks are due Lynne D. Dubiner, Project Secretary, for her competent assistance in administration of the grant, for typing and proofing of the several manuscripts, and for her unfailing cheerfulness and interest. The period of association with the grant is indicated for each member of the project in Appendix C.

It is gratifying to report the widespread interest of the medical center faculty in the research project. An important example of this interest is the fact that six members of the faculty who were not formally associated with the research staff share authorship of one or more of the studies. They are June C. Abbey, R.N., M.S., Visiting Assistant Professor, School of Nursing; R. L. Blancheri, D.D.S., Associate Professor and Chairman, Crown and Bridge Section, Division of Prosthodontics, School of Dentistry; Velena Boyd, R.N., M.P.H., Assistant Professor, School of Nursing; Edith S. Deck, R.N., M.S., Associate in Psychiatric Nursing, School of Nursing; S. F. Lorencki, D.D.S., Assistant Clinical Professor, Crown and Bridge Section, Division of Prosthodontics, School of Dentistry; and N. C. Staub, M.D., Assistant Professor, Department of Physiology and member of the Cardiovascular Research Institute, School of Medicine. The valued contribution of each author is indicated as a footnote at the beginning of each chapter.

Grateful acknowledgement of assistance in preparation and execution of one or more research studies is due W. F. Ganong, M.D., Patricia A. Hoyt, R.N., Alice E. Ingmire, Ed.D., R. H. Kellogg, M.D., F. Anne Pirnie, R.N., M.A., and Ann W. Roweder, R.N., M.S., all of the University of California Medical Center, and Marylou McAthie, R.N., M.S., San Joaquin General Hospital, Stockton, William McFadden, R.N., Fresno Community Hospital, Fresno, and Alice Pelletaire, R.N., Santa Clara County Hospital, San Jose.

Recognition of the valuable contribution of broadcasts in the public interest is due

television stations KJEO, Fresno, KNTV, San Jose, and KVIE, Sacramento. Other organizations which graciously participated in one of the studies include Mt. Zion Hospital, San Francisco, and Fresno County Hospital, Fresno.

The authors of the several studies are particularly grateful for the valuable consultation of Mary L. Epling, M.A., Statistician, Cancer Research Institute, and Anne M. Schmid, Editor, Department of Pediatrics.

When the several studies were completed, findings were submitted to appropriate journals in specialized fields (Grant, 1959, p. 7). Material of this final report has been published or is in press, as follows:

Chapter I and Appendix A--J. med. Educ., 1963, 38, 329-338.

Chapter III--J. med. Educ., In press.

Chapter IV--J. dent. Educ., 1962, 26, 146-151.

Chapter V--J. dent. Educ., 1963, 27, 167-170.

TABLE OF CONTENTS

Page

I. TELEVISION CAPABILITIES AND INSTRUCTIONAL EFFICIENCY 1

II. HOME AND HOSPITAL VIEWING OF CONTINUING EDUCATION BROADCASTS UNDER THREE PRESENTATION-RESPONSE CONDITIONS 8

III. MEASURING ACHIEVEMENT IN SCIENCE REASONING ABOUT PHYSIOLOGY . . . 22

IV. EFFECTIVENESS OF TELEVISION WITHIN THE DENTAL LABORATORY 28

V. THE STEP PRESENTATION OF DENTAL TECHNIC INSTRUCTION 34

VI. CAMERA PLACEMENT FOR RECOGNITION OF COMPLEX BEHAVIORS 38

APPENDIXES

A. INSTRUCTOR - TELEVISION STAFF COOPERATION 45

B. PERCEPTUAL RECOGNITION OF PICTORIAL SIGNS 46

C. RESEARCH STAFF 53

REFERENCES 54

I. TELEVISION CAPABILITIES AND INSTRUCTIONAL EFFICIENCY*

A recent analysis of the findings of 400 closed-circuit television studies classified them according to subject matter of instruction. Television proved more efficient than conventional face-to-face instruction in 29 per cent of the studies of health and safety and 25 per cent of the studies of science instruction. This compares with findings favoring television in studies of language or humanities, history, and art of only 17 or 11 per cent, respectively (Schramm, 1962).

There appears to be good reason why televised instruction has proved efficient in the area of the health sciences: the subject matter is particularly suited to the capabilities of the television system. Certain of these capabilities already have been identified. Additional research can increase the efficiency of television.

Television Capabilities

The following discussion leans in part upon earlier statements by Ruhe and Proud and by Kanner (Ruhe & Proud, 1958; Kanner, 1958). The efficiency of television in the communication of health sciences education appears related to its capability of image multiplication, magnification, association, transportation, transformation, and memory.

Image Multiplication

Image multiplication is the simultaneous display of the same information at two or more physically separated locations. This capability is usually applied to mass instruction through lecture or panel discussion. A typical application occurs in the continuing education of the health sciences practitioner. For example, the largest auditorium at the University of California Medical Center in San Francisco has a seating capacity of 438, yet the installation of television viewing monitors at three other locations in the Center made it possible to accommodate a total of 1,081 participants at a recent seminar entitled, "Man and Civilization: Control of the Mind" (Merrill, 1962).

When compared with a face-to-face lecture, image multiplication does not appear either to interfere with learning or to enhance it. Tannenbaum, who divided medical students in a physiology course into a television viewing group, a face-to-face group, and a television originating-room group, could find no statistically significant difference between the groups as measured by an achievement test (Tannenbaum, 1956). Macomber and Siegel investigated performance of undergraduate college students in each of 12 courses by dividing them into a conventional class-size group, a large class group, and a television group. No statistically significant differences were found on measures of achievement, problem-solving and synthesis, or stereotypes and misconceptions (Macomber & Siegel, 1960).

Whatever gain in efficiency accrues from image multiplication must be measured in terms of cost. When Carpenter and Greenhill compared television with the cost of conventional instruction at Pennsylvania State University they found television saved money when 200 or more students could be taught a given course at a given time (Carpenter & Greenhill, 1958). No similar study for medical education has been reported.

Image Magnification

Magnification is display of a television image with enlargement of plane dimensions to a minimum of 50 times those of the actual object, and can approach 100 times by

*Irving R. Merrill

means of the television camera lenses alone. Higher powers can be achieved by means of supplementary lens systems, such as microscopes. The maximum power of a given microscope may be multiplied an additional 50 to 100 times by the use of the television system.

A typical example is found in histology instruction. Each student enters the laboratory and sets up his individual microscope. The instructor then presents slide specimens similar to those with which the student will subsequently work. Because the television system can be adjusted in advance, the instructor merely changes slides and adjusts the microscope to direct attention to various portions of the specimen by means of a separate instructor-monitor. Following this television exposition, the student examines his own set of slides. If the laboratory assistants report that a number of students experience difficulty during any particular part of the exercise, the instructor uses television to clarify the problem.

Magnification is also employed in pharmacology. Without a supplementary lens system the effect of certain agents on salivation of an experimental animal can be demonstrated so that each student in the class can count the number of drops of saliva leaving the end of the canula at various stages of an experiment. Each drop becomes 6 inches wide when displayed on the viewing monitor.

There are no reported studies of instructional efficiency based upon the magnification capability.

Image Association

Association is the simultaneous display on a single viewing monitor of the signal from two or more live cameras. This may be accomplished by superimposition or by the "split screen" effect obtained from a special effects amplifier. This capability is widely applied to the demonstration experiment in the basic health sciences.

One problem in physiology instruction is that of helping the student to associate such functions as blood pressure, inspiration, and electrical activity of the heart and brain with their graphic representations. This is difficult even when three or four students attempt to observe a laboratory demonstration; even though the recording pen is located beside the table, a constant shift of vision between the animal and the pen is necessary. With the split screen device, these images can be placed adjacent to each other and observed by any number of viewers.

The split screen technique has also been used in the clinical areas. A recent application to psychiatry involved the interview. The patient and interviewer faced each other. At a considerable distance behind each individual was a camera equipped with a long lens to obtain a full-face view of the other person. Each image was then displayed side-by-side on the same viewing monitor so that both persons appeared to be facing the viewer. The subtle interplay of facial expression, an important part of non-verbal communication, was thus added to the simultaneous unfolding of spoken dialogue. Although the split screen method of association appears to be most desirable in each of the above examples, the superimposition technique has been substituted with success.

While the split screen technique associates images by placing them side-by-side, super imposition accomplishes association by laying one picture on top of another. By use of the switcher-fader, either camera signal may be made more or less prominent, or faded-out completely. An almost perfect application to health sciences instruction was suggested recently to the writer by W. O. Reinhardt, M.D., Chairman, Department of Anatomy. One objective of the study of anatomy is proficiency in relating external configuration of the body to the skeletal structure of, for example, the chest. This relation may be dramatically demonstrated by television. The instructor indicates the

significance of the external configuration, draws the underlying structure at the presumed location on the skin, then, through superimposition of a previously completed chest X-ray, the rib cage appears within the chest to verify the accuracy of the marks on the body.

There are no reported studies of instructional efficiency based upon the association capability.

Image Transportation

Image transportation is the movement of an image from or to an otherwise inaccessible location for purposes of display. In the health sciences accessibility is a particularly acute educational problem.

Certain locations are rendered inaccessible for reasons of physical safety. In areas of radiation hazard, the safety of the student is of primary concern; within the operating room, the safety of the patient requires his isolation from direct contact with student observers. The use of television for this type of image transportation is too well known to require documentation.

Less obvious, but potentially equally important, is the transportation of an image into an otherwise inaccessible location. For example, instructions on laboratory procedures are best presented immediately before application. Students are least likely to forget under such conditions, and the immediate application of the procedure tends to fix the correct performance more firmly in the mind of the student. In many laboratories, room configuration and objects within the room preclude delivery of these instructions as an illustrated lecture complete with screen, projector, and photographic slides. Television can overcome this inaccessibility by transporting the image directly to the laboratory table of the student.

In a recent experiment instructions on a procedure to be performed in the laboratory were presented as a half hour demonstration to a group of students in the classroom and to a similar group in the laboratory. Both groups were given a fixed interval of time to complete a test. The laboratory group, which applied immediately what had just been learned, performed significantly better than the conventional classroom group who had not gone directly to the laboratory (Grant, Biancheri, Lorencki, & Merrill, 1962).

In addition to absolute inaccessibility due to hazardous conditions, other areas may be relatively inaccessible to students because of distance, time, or expense. The ability of television to transport an image across these often formidable obstacles is receiving increasing consideration. It may prove more efficient to transport a certain amount of clinical instruction from a teaching hospital in a large city to students in a two-year medical school in a small college town by television than to transport the students to the hospital. The visits to the teaching hospital may then be devoted to those learning experiences which can be obtained only at the bedside in person (Hard, 1962).

Image Transformation

Image transformation is the display of attributes of an object which would not be visible without special alteration of the electronic components of the television originating subsystem, and does not depend upon magnification. Although it has high potential for increasing the efficiency of instruction, this capability does not appear to have been employed as frequently as those already described.

Transformation can be accomplished in several ways, one of which is through change in the television camera tube. Zworykin has reported development of an experimental vidicon tube which is sensitive to ultraviolet light so that substances which appear

transparent when exposed to white light can now be revealed clearly (Zworykin, 1957). Because ultraviolet rays are not visible to the human eye, this capability of television is particularly significant.

Within the visible spectrum, the normal television system will display tones of light with a contrast range of 30 to 1. The human eye, however, is capable of seeing at least a 100 to 1 contrast range in an actual scene. The television system is capable of reducing a greater contrast range to its own system limits, but it is also capable of expanding a smaller contrast range into a 30 to 1 ratio. This type of image transformation is applied frequently to X-ray study, in which differences in tone too minute to be visible to the naked eye may be seen clearly on the viewing monitor. If a color system is employed in this fashion, a further contrast is achieved which shows relationships of spatially separated areas (Fisher & Gershon-Cohen, 1957).

This "clipping" of a portion of the black and white television range to brighten the remaining portion electronically need not be limited to X-ray materials. It can be used, for example, to show a solid object within an apparently opaque solution.

Image Memory

Image memory is the subsequent display of information originally obtained from a live television camera, and is finding increasing application to health sciences education.

One type of memory permits the instructor to interrupt a sequence of action and display a single frame for a given period. Television normally displays a new picture of 525 lines every 1/30 second, and operates at 30 frames per second, just as sound motion picture projection operates at 24 frames per second. This type of television memory is similar to a motion picture sequence held in one frame on the screen for 10 minutes for careful study. A typical application is in television fluoroscopy which permits a number of faculty members and students to make an immediate and detailed study of a case and eliminates the time required for development of the film. This type of memory is accomplished by storage monitors, storage tubes, or special vidicon camera tubes. The duration of a single frame display can be controlled by the operator by switching off the screen and instructing the device to store the frame for a limited period for subsequent display. When all the power supplied to the device is switched off, the frame cannot be retrieved.

Another type of memory permits the instructor to record an entire sequence on a more permanent material and retrieve it for display at any subsequent time. The first device developed for this purpose was the kinescope recorder, which produced a 16 mm sound motion picture from the face of a kinescope (viewing) tube. More recently, the videotape recorder made it possible to lay down the electrical impulses on magnetic tape for immediate playback if desired.

Libraries of television recordings of value to medical education are known to exist at the University of Michigan (color and monochrome), Walter Reed Army Medical Center (color and monochrome), the National Naval Medical Center (monochrome), the University of Utah (monochrome), and the University of California Medical Center, San Francisco (monochrome).

No comparisons of the efficiency of live to recorded television in health sciences

education have been reported. The most recent study involved military training where it was found that Army personnel learned as well from a recording as from live television (Kanner, Mindak, & Katz, 1958). Television recordings were enjoyed as a novelty over motion picture film at first, but as early as 1955, Hurst found that student performance was not affected whether films or kinescopes were announced as kinescope recordings, or the same materials were announced as films (Hurst, 1955).

Availability of the memory capability adds much to the flexibility of the television system. For example, the large attendance at the symposium on "Man and Civilization: Control of the Mind" indicated that the sessions would also be of interest to the general public, and a local television station requested recording for subsequent broadcast. Thus, the closed-circuit system made simultaneous broadcast recording possible as well as on-campus multiplication (Merrill, 1962). Because an international faculty had been assembled for the symposium, this instance illustrates what appears to be a general principle on application of this memory capability: to record those television presentations which can never be repeated live or which can be repeated only with great difficulty.

So much for the evidence in support of the hypothesis that these six instructional capabilities are the reason why television has proved efficient in health sciences education. From its beginnings in the laboratory to its culmination in direct personal contact with the patient, the visual element is one of the distinguishing characteristics of health sciences instruction. The capability of television to serve instructor and student by multiplication, magnification, association, transportation, transformation, and memory of this element may account for its success.

Instructional Efficiency

Several lines of research can result in further improvement of the efficiency of televised instruction. The most immediate results may be expected from continuation of several programs of applied research on health sciences instruction. A greater potential gain counterbalanced with a somewhat lower probability of success might be found in basic research on the learning process through the use of television as a research instrument. A final possibility lies in electronic research, in the design and discovery of radically different television systems.

In applied research, immediate attention should be given to linking system capabilities to the principles of learning. Large increases in efficiency can thus be obtained. For example, a gain of moderate practical importance resulted when transportation of a conventional length instructor-demonstration was compared with classroom viewing (Grant, et al., 1962). A large additional gain in efficiency occurred when the transported presentation was arranged to conform to the memory span of the students (Blancheri & Merrill, 1963).

Applied research studies may result in the discovery and classification of presently unrecognized television capabilities which in turn should be investigated. In addition to the grant to the University of California Medical Center, San Francisco, the U.S. Office of Education has provided support for two other experimental studies within the health sciences, which are recently completed or in progress. Project 074, Baylor Medical College, Houston, W. A. Spencer, Principal Investigator, concerns teaching human physiology via a multi-channel data broadcast system. Project 322, New York University College of Dentistry, New York City, A. H. Morrison, Principal Investigator, is an experimental study utilizing closed-circuit television in the teaching of dental techniques.

Another area of applied research lies in the study of how television characteristics affect learning. A case in point is a recent study of camera placement based on a prior

study of Roshal, who compared the efficiency of two camera angles in teaching a perceptual-motor task to Naval trainees. In the "subjective" camera angle, the students saw what the instructor saw as he demonstrated the correct procedure; in the "observer" camera angle, the students saw what an observer seated opposite the instructor would have seen. Roshal found the subjective angle superior (Roshal, 1949). A more complex task was the instructional objective of the present study -- i.e., what a nursing team leader does in a daily nursing care conference. Nurses who viewed the observer angle version of the demonstration scored higher on a recognition test than the group who viewed the subjective angle version (See Chap. VI, Eds.). It is clear that there is still much to be learned about what determines the point of optimum perception.

There are good reasons for encouraging the occasional use of a health sciences closed-circuit television system as a research instrument for the study of human learning. Faculty in the departments of medical psychology, psychiatry, physiology, ophthalmology, and otology have specialized interests in this general problem and possess the background of knowledge required for the solution of pertinent studies in this area.

Television possesses many advantages as a research instrument. The multiplication capability permits the simultaneous display of identical information in separate geographical locations. The memory capability permits the advance preparation of experimental subject matter, so that its nature may be determined exactly, and its successful presentation during the experiment assured. Many learning experiments involving visual display of subject matter to one student at a time could be expedited by image multiplication. Tachistoscopic devices could be integrated into the system. Multi-sensory experiments, including auditory only, auditory-plus-visual, and visual stimuli only, can be facilitated by television. Findings from such studies will benefit health sciences education and general education as well.

Health sciences instruction may expect to gain from experiments in human learning and in the psychology of perception which are (most appropriately) performed elsewhere; improvements in the basic design of the television system also fall in this category.

There are some who are disturbed by talk of efficiency in medical education. To them, the term often implies impersonal mass production of a uniformly mediocre product -- cheap education. This view of efficiency is not accurate in two respects: (a) it confuses the goals of instruction with a system of measurement; and (b) it assumes that efficiency can only be measured in terms of money.

The medical educator is free to state his own goals. If they are seen merely in terms of a limited set of items of specific information, so much the worse for us all. He is also free to specify other objectives for his students, such as critical thinking, science reasoning, diagnostic excellence and creativity. Whatever the educational goals, if the instructor has alternative means of achieving them, he has an obligation to consider the relative efficiency of those alternatives.

There are a number of different measures of efficiency employed today, and cost is only one of them. Others, such as time or effort, may be far more relevant in certain situations. One may hold time constant and measure the percentage of task completed (During the freshman year, can science reasoning ability be improved more by television or by conventional instruction?). One may specify the task to be completed and measure time required to complete the specified task (Can mastery of a laboratory procedure be taught in less time by the step presentation of televised instruction?). It is the responsibility of the investigator to make explicit the measure of efficiency he will use and demonstrate its suitability to the problem at hand (Ackoff, 1953).

After only a few years the closed-circuit television system has earned a permanent place in the health sciences curriculum. Continued, vigorous exploration of its potential

through research and on-going instruction can make closed-circuit television even more efficient in communicating the recent extraordinary proliferation of knowledge within the health sciences.

II. HOME AND HOSPITAL VIEWING OF CONTINUING EDUCATION BROADCASTS UNDER THREE PRESENTATION-RESPONSE CONDITIONS*

This study compares the instructional efficiency of home and hospital viewing of continuing education broadcasts under three presentation-response conditions. There is no longer any question as to whether open-circuit television broadcasts can transmit instruction effectively to professionals within the health sciences, and the critical question for the use of television is how to use it. Since there is little evidence about the most efficacious ways to use television, this study examines two critical variables relating to instructional efficiency:

1) the conditions of viewing -- Will a physician or nurse gain more from a broadcast which is viewed individually at home or with a group of colleagues in a hospital conference room?

2) the degree to which student participation in the broadcast should be encouraged -- What conditions of participation produce increased learning when compared with the conventional broadcast which does not require active participation?

Review of Pertinent Literature

A review of the literature covers three problem areas: (1) what are the most effective conditions for viewing, (2) how much does student participation help learning, and (3) what are the most important variables related to attendance throughout the broadcast course.

Evidence on the conditions of viewing is inconclusive, but it favors individual viewing at home. Lepore and Wilson and Dreher and Beatty compared achievement of freshman and sophomore college students who had been divided into groups for individual viewing at home, group viewing on campus, and face-to-face instruction. No consistent differences between groups appeared over the several courses studied (Lepore & Wilson, 1958; Dreher & Beatty, 1958). A similar comparison was made of groups of junior college students by Erickson and Chausow. Individual viewing at home was found highly superior to group viewing on campus for a course in social science, but no differences among groups were found for courses in physical science and speech (Erickson & Chausow, 1960). For the social science course the superiority of home viewing may have been caused by differences in group composition, rather than by the conditions of home viewing.

Student participation is not entirely absent from the conventional presentation which expects to keep the viewer alert and responsive through effective organization of subject matter to attain a clearly defined objective. One way of increasing participation is to insert factual questions into the conventional presentation. Students may make an overt response by writing the answers, or they may make a covert response upon being told to think of the answers for themselves. Michael, who compared achievement of several groups of high school students after viewing a film, found that groups encouraged to make a covert response to inserted participation questions scored higher than those who viewed the conventional film presentation (Michael, 1951).

The major gain in achievement from inserting participation questions was found to occur on test items directly covered by the questions (Lumsdaine, May, & Hadsell, 1958). It has also been determined that the greatest gain from active student response to participation questions occurs when the subject matter is most difficult (Hovland, Lumsdaine, & Sheffield, 1949). Additional gains in achievement occurred from participation questions if the student was given immediate knowledge of results or feedback, i.e., if the student was informed of the correct answer immediately after he responded to the

*June C. Abbey, Velena Boyd, Edith S. Deck, John G. Shrock, and Irving R. Merrill

question and before the presentation continued. Michael found feedback was the greatest contributor to learning when compared with the variables of overt vs. covert response and informing the students prior to viewing that they would be tested afterwards (Michael, 1951). Junior high school science students who saw a programmed television lesson which required active student response scored higher than students who saw a conventional lesson on the same subject matter, when tested immediately after instruction. When retention was measured two and a half weeks after instruction, there was no difference between the scores (Gropper & Lumsdaine, 1961). An excellent review of active participation by Allen cites additional studies in which these findings are supported (Allen, 1957).

This review of the variables that increase learning indicates that individual viewing at home may be preferable to group viewing, but that more conclusive evidence is needed. It also indicates that encouragement of student response to participation questions inserted into an effective conventional presentation produces gains in learning. Moreover, these gains can be enhanced by providing the student with immediate knowledge of results, or feedback. Little is known as to whether these gains over the conventional presentation are retained for any length of time or whether they apply to individuals viewing at home.

Attendance is important if television is to be used most efficiently. Dreher and Beatty noted that for televised college courses in psychology, economics and basic communication, the percentage of students who failed to complete or who withdrew from the courses ranged from 3-16 per cent for those who viewed as a group on campus and from 9-28 per cent for those who viewed individually at home (Dreher & Beatty, 1958, p. 44). Similar results were reported in junior college courses by Erickson and Chausow (1960). In the series of open-circuit teleclinics for physicians on diabetes mellitus, viewing declined as the series progressed. An average of about 10 per cent of the sample of physicians contacted by telephone had watched any given telecast in the series. Calls were made by a professional survey organization within two days of a given telecast. For physicians responding to a mail questionnaire, this figure was about 18 per cent (A Staff Summary and Critique of a Study on the Utah Open-Circuit Teleclinics for Practitioners, 1961).

The general problem of this study was to apply these theoretical findings to the practical problem of how television can best bring continuing education to those located great distances from any medical center. The application of conditions of viewing was straightforward. However, the most effective presentation-response condition was thought to be the one which also does most to encourage the professional to view. One condition selected was the broadcast of a conventional presentation. Another condition was the broadcast of the same presentation with participation questions inserted and with the students encouraged to think of the answers for themselves. A third condition rewarded the student for making an overt response to the participation questions by providing him knowledge of results by return mail if he forwarded his completed answer sheet to the medical center. He could only obtain the questions if he viewed the broadcast.

The specific purpose of the study was to test the following hypotheses: I. The learning achievement and regularity of attendance are greater for general hospital staff nurses who view a continuing education broadcast individually at home than for those who view the same broadcast as a group in a hospital. II. The learning achievement and regularity of attendance vary in descending order for the following presentation-response conditions:

a) overt response to participation questions with feedback,

b) covert response to participation questions without feedback, and

c) covert response to a conventional broadcast without feedback.

Method

This experiment involved the cooperation of three VHF television stations, six hospitals, and over 150 active professional nurses. Because of its complexity, a first phase, the pilot study, was conducted locally by use of kinescope recordings. The second phase, the main study, was conducted by broadcast television in areas located 50 miles or more from the medical center.

Pilot Study

The purpose of the pilot study was to test the hypothesis that learning achievement under the overt presentation-response condition exceeds achievement under the conventional conditions, as measured by an external examination following the presentation.

The Subjects (Ss) were 35 local general hospital staff nurses who were paid to participate in the experiment; they were randomly assigned to four groups with nine Ss in Groups I, II, and III, and eight in Group IV. Groups I and III viewed under the overt condition and Groups II and IV viewed under the conventional conditions. Groups I and II were tested one week before the others.

The following procedure was used on both days: Both groups were assembled for a brief orientation in Room A at which time group assignments were announced. One group then moved to Room B after which Ss in both groups received an envelope with reading material about the subject matter. The overt condition group also received written instructions on how to respond to the participation questions plus an answer sheet for their written responses. Each group then viewed the first of two kinescope recordings under the appropriate condition. Time was allowed for the answer sheets in the overt condition group to be collected, checked, and returned with an explanation sheet. The second recording was then presented, with the answer sheets again handled in the same way. Both groups were then reassembled and asked to respond for half an hour to some questions which were not related to the subject matter, after which both groups received a 20-item examination. Ss were allowed to leave with instructions not to discuss the experiment with anyone.

Room assignments for the overt and conventional condition were rotated on the second week to balance out possible room differences such as acoustics. The interpolated questions were included to determine if differences related to presentation-response conditions were retained for at least half an hour. A much longer interval between viewing and testing was planned for the main study.

The stimulus materials consisted of two programs on "Emphysema" scheduled for use in the main study, and the pilot examination consisted of 20 items scheduled to be among the 60 items in the main study examinations. Ten of the pilot test items were related to corresponding participation questions in the programs on emphysema, while the other 10 test items sampled learning of subject matter not covered by the participation questions. Materials and examination are discussed more fully in connection with the main study.

Main Study

The experimental design was a three-by-three Graeco-Latin square, in which groups in three geographically separated cities viewed television programs on three different topics of nursing in three different orders under three presentation-response conditions. Within each city some Ss viewed individually at home and others viewed as a group at the hospital.

The stimulus material was a series of six programs entitled Recent Advances in

Nursing Care of the Aged with two programs devoted to each of three topics: "Emotional Illness," "Emphysema," and "Cerebral-Vascular Accidents." Each topic represented instruction from a major area of nursing, i.e., psychiatric, medical-surgical, or public health nursing. Because of these differences in subject matter, the television production techniques varied among the topics. Each topic was prepared and presented by a different Experimenter (E). Thus, if significant differences were observed among the experimental conditions, the differences would be known to be stable for at least three separate topic-teacher-television technique situations.

The series was designed for the instruction of a specific member of the nursing profession: the staff nurse in a general hospital. The material on each topic was new to most viewers.

The broadcast version prepared for the overt and covert conditions was 28 minutes in length, including 10 participation questions, which appeared at about three-minute intervals during the broadcast. The questions appeared visually on the screen for approximately 30 seconds. After 10 seconds the instructor spoke the question, leaving time for the S to write or think the required one-word answer. The questions were omitted from the conventional condition, reducing the length of the presentation to 23 minutes. A set of five-minute films prepared by the American Red Cross on "Ice Rescue" and "Tornado Disaster Relief" were supplied to each television station to fill the broadcast time with non-related material during presentations under the conventional condition.

The Ss were 120 staff nurses, 40 from each of the three cities, currently employed in a general hospital. To make the experimental situation more closely resemble a regular continuing education course, the experimental series was sponsored by the School of Nursing as a course which carried a certificate of completion. Each S was a volunteer who, through an enrollment fee of $3.00, paid to participate in the experiment. Ss were informed that the course was experimental in nature.

Because it had not been possible to obtain equated groups for home and classroom viewing in some of the studies reviewed, careful attention to this matter was given here. Within each city 20 Ss were assigned to individual viewing at home and 20 to group viewing at the hospital. These groups were equated on age, number of children -- a possible distraction for home viewing, and hospital shift. When data for all cities were combined, the 60 Ss assigned to home viewing and the 60 assigned to group viewing in the hospital compared as follows:

	Individual Viewers	Group Viewers
Mean age (age range)	43.26 yr. (24-61)	43.33 yr. (21-61)
Mean number children	1.1	1.1
Day shift	48	52
Afternoon shift	4	7
Night shift	9	1

The cities selected for the study were widely separated from the medical center and from each other. Ss in one city could receive a strong signal from the television station serving that area but none from the two stations carrying the broadcasts for the other cities. Two stations used videotape and one used kinescope film for the series. In city A the time of viewing was Tuesdays at 1:30 p.m. Because all 20 Ss assigned to group viewing in this city could not meet together, 12 viewed together at one hospital and 8 viewed at another comparable facility. In both cities B and C, group viewing was done at one location at 4:00 p.m. Tuesdays and 4:30 p.m. Wednesdays, respectively.

Learning achievement was measured by three 20-item topic tests, accounting for a total of 60 multiple choice items for the series. The S's score on each test was the number of items answered correctly. Learning from each program was sampled by 10 of the test items. Each topic test was assembled on the basis of an item analysis of a pre-test given to a group comparable to the experimental Ss. The degree of relationship between test items and the participation questions which appeared on the television screen under the overt and covert conditions is shown by the following example:

Participation Question: Supervision of the emphysematous patient with hallucinations should be _____. (Answer: continuous.)

Related Test Item: Emphysematous patients with hallucinations need mainly which one of the following in nursing care:

A. Discontinuation of ACTH

B. Mild medications to calm the patient

C. Many reassurances about reality

D. Continuous supervision

E. Restraints

(Answer: D)

Measurement of attendance was necessary to obtain a comprehensive analysis of the functions of the two experimental variables. Information about the effect of the conditions of viewing and the presentation-response conditions upon regularity of attendance were important objectives of the study. Attendance was measured in two ways. Reported attendance was obtained by having every S who completed the course indicate on a special form which of the programs had been viewed. Actual attendance of the Ss at the hospital was monitored at each broadcast in each city by members of the research staff. Copies of the learning and attendance instruments and videotape copies of the broadcast series are available from the Office of Television Research.

The series was broadcast for six weeks beginning April 16, 1963, and an evaluation session was conducted in each city on the seventh week. The experimental procedure included the use of reminder information, control of presentation-response conditions, and conduct of the evaluation session. Reminder information consisted of a brochure, program announcements, and an evaluation session letter. The brochure, which was distributed to all Ss prior to April 16, described the course, and indicated for each city the title and broadcast date of every program. Prior to each program all Ss received an announcement through the mail of the program for that week. The evaluation session letter indicated time, date and place of the session and carried a reminder that award of the certificate of completion depended upon attendance in person, including those who viewed individually at home.

For the conventional condition the S received only the program announcement. Under the covert condition, the S also received instructions on how to respond. Under the overt condition the S received an answer sheet and instructions to write the answer to each of the participation questions as they appeared. The questions were not given on the answer sheet, since it was necessary to view the broadcast in order to obtain the questions. A research staff member collected the completed answer sheets at each hospital immediately after presentation. Home viewers mailed their answers in a stamped addressed envelope to the medical center research office. From three to five days after the presentation, the S received by mail his own answer sheet without a grade

and an explanation sheet which contained question and correct answer with an explanation. Ss who failed to return an answer sheet did not receive the reinforcement of the explanation sheet.

At the evaluation session in each city the Ss first completed the reported attendance form. They were asked to be accurate and were assured award of the certificate of completion did not depend upon high attendance. This report was obtained first to prevent any response bias resulting from the S's estimate of his own performance on the three topic tests. The Ss then completed the three topic tests, which were stapled together in the order in which the topics had been presented in that city. Ss were told scores would be held in confidence and would not affect award of the certificate. Ss were invited to remain for an open-ended discussion of the series and the experiment. At the conclusion, Ss were requested not to discuss the evaluation session, since a make-up session was to be held the following week for Ss not able to attend on schedule.

Results

Pilot Study

To test the hypothesis that learning achievement under the overt presentation-response condition exceeds achievement under the conventional presentation-response condition, as measured by an external examination following the presentation, the means of the two conditions were compared. Although the overall mean for the overt presentation-response conditions exceeded the mean for the conventional presentation-response conditions, the difference was not significant.

A more detailed analysis of the data was performed to answer the specific question of whether or not the participation questions increased learning. For the two experimental groupings the responses to the 10 test items that were related to corresponding participation questions were compared to the responses to the 10 items that were not covered by participation questions. The mean learning achievement score for the 18 Ss in Groups I and III (overt condition) was 7.56 for the 10 test items related to participation questions and 5.76 for the 10 non-related items. For the 17 Ss on Groups II and IV (conventional condition) the respective means were 6.35 and 6.47. In order to examine the interaction between achievement scores for related and unrelated test items for the presentation-response conditions, a Chi-square test was performed on the data (X^2 = 6.2 with 1 df), which was significant between the .02 and .01 levels of significance. The data are presented in Table 1 in which the columns represent the two presentation-response conditions, and the rows divide the Ss according to whether they received a higher score on the ten related test items, the 10 unrelated items, or received the same score on both sets.

Table 1

Number of Subjects For Two Presentation-Response Conditions
By Score Comparison on 10 Test Items Related and 10 Items
Not Related to Participation Questions

Test Item Scores	Overt	Conventional
Related Score Greater Than Unrelated	12	5
Both Scores Equal	4	2
Unrelated Score Greater Than Related	2	10
Total Subjects	18	17

With ties not considered, $X_y^2 = 6.2$; $X_{.98}^2 = 5.4$ for $n = 1$

Under the conventional condition, which contained no participation questions, the groups scored numerically lower on the related test items. Although the mean difference was not significant, the data suggests that the related items were slightly more difficult than the unrelated items. In contrast to the conventional condition the overt presentation-response condition was superior on the more difficult related test items. This parallels the findings of Lumsdaine, May, and Hadsell (1958), and Hovland, Lumsdaine, and Sheffield (1949).

The pilot study findings were helpful in the preparation of broadcasts and tests for the topics "Emotional Illness" and "Cerebral-Vascular Accidents." Because the data indicated that the 10 related participation questions could be effective in increasing learning of health sciences subject matter for a group which viewed at a hospital, this format was followed throughout the series. Test items for these topics were then related as carefully as possible to the appropriate participation questions, as in the example presented in the section on Method.

Main Study

Attendance. Of the 120 Ss who paid to take part in this experiment in continuing education, 109 participated in the evaluation session and received a certificate of course completion. For the 60 in the three cities who viewed with a group in a hospital, the percentage who failed to complete the course was 9.2. For the 60 who viewed individually at home, the percentage was 9.0. It is clear from inspection that these two groups did not differ in this respect. These records are better than the rates per semester course reported by Dreber and Beatty for undergraduate students studying under similar conditions of viewing (Dreber & Beatty, 1958).

The following analyses are based upon the 109 Ss who completed the course. Attendance was very high throughout the series. It was possible to compare the average rate of actual and reported attendance for the 54 Ss who viewed as a group in a hospital. An average of 79 per cent of these Ss actually attended each broadcast. Based upon reported attendance, an average of 90 per cent of the Ss attended each broadcast. A tendency to overreport resulted in an increase of 11 per cent over actual attendance, which is similar to the percentage difference between the telephone survey (10%) and the reported attendance (18%) for physicians, who viewed the series on diabetes mellitus (A Staff Summary and Critique, 1961). Elements which may account for the high rate of attendance in this study are enumerated in the Discussion.

Regularity of attendance figures provided information on whether either experimental variable acted to prevent a decline in attendance as the series progressed. Regularity of attendance was defined as the proportion of Ss present for the first program who also were present for the second program of a given topic. This proportion indicated the pulling-power within a given experimental condition independent of the effect of the previous events during the series of broadcasts.

The formula $\phi = 2 \arcsin \sqrt{P}$ was used to transform each proportion into a value which could be used in the analysis. This formula is valid with only slight inaccuracy when the proportion falls within the ranbe .05 - .95 and when N is 20 or more (Walker & Lev, 1953). The restriction on the size of the proportions eliminated the possibility of using reported attendance data, because too many proportions were greater than .95. It was originally planned to use reported attendance in this analysis because it was available for all 109 Ss. Regularity of attendance proportions in Table 2 were prepared from data on actual attendance. This limited the analysis to Ss who viewed as a group in a hospital.

An analysis of variance, following the conventional Graeco-Latin Square design, was performed on the nine transformed proportions; the error term was corrected for

unequal numbers. In the analysis one proportion exceeded .95 and the numbers of Ss per city were 18, 20, and 16. The four main effects in Table 3 are (A) Presentation-Response conditions, i.e., Overt, Covert, and Conventional; (B) Order, i.e., appearance First, Second, or Third in the series; (C) Topic, i.e., "Emotional Illness," "Emphysema," or "Cerebral-Vascular Accidents"; and (D) City, i.e., City A, B, C. None of these produced statistically significant differences in regularity of attendance. The overt condition was no more efficient than the covert or conventional conditions. The non-significant F-ratio for (B) indicates that attendance did not decline as the series progressed; attendance was the same throughout. Only (C) approached statistical significance, which suggested that attendance was more regular for one topic than the others, but even this difference could be accounted for by chance alone. (D) indicates that Ss in all cities attended with similar regularity.

Table 2

Proportion of Those Actually Viewing First Program in Hospital Who Also Viewed Second Program by Three Presentation-Response Conditions

	Overt	Covert	Conventional
City A	1.00	.88	.79
City B	.88	.94	.78
City C	.71	.93	.86

The first hypothesis (conditions of viewing) was not tested in the analysis, because the actual data applied only to Ss viewing as a group in a hospital. As a result, there was no evidence whether regularity of attendance was greater for the general hospital staff nurses who viewed individually at home. Since the reported attendance was highly similar for Ss under both conditions of viewing, it seems likely that the regularity of their actual attendance did not differ.

The second hypothesis (presentation-response conditions) was not supported by the analysis of regularity of actual attendance. Since regularity of attendance revealed no differences for the experimental variables, any differences in learning achievement could not be attributed to attendance.

Table 3

Analysis of Variance in Regularity of Actual Attendance

Source	Sum Sq.	d.f.	Mean Sq.	F.	Significance
(A) Presentation-Response conditions	.17	2	.085	1.16	
(B) Order	.17	2	.085	1.16	
(C) Topic	.38	2	.190	2.60	a
(D) City	.09	2	.045	.62	
Total	.81	8			
Error		∞	.073		

[a]Exceeds F.90 = 2.30 at 2/∞ d.f.

Learning achievement. The analysis of learning achievement was based upon the topic test scores of each S. No transformation of scores was required, but some adjustment of score was necessary for two topics, because in City B the broadcast sequence of two programs was reversed. Instead of the specified order "Emotional Illness, I and II (covert)," followed by "Emphysema, I and II (overt)," Ss received the four programs in the following order: (a) "Emotional Illness, I (covert)," (b) "Emphysema, I (covert)," (c) "Emotional Illness, II (overt)," and (d) "Emphysema, II (overt)". This had the effect of confounding topics but not presentation-response conditions, for one city. However, programs (a) and (d) were broadcast as specified.

No adjustment was necessary for topic scores on "Cerebral-Vascular Accidents". For "Emotional Illness" the S's score on the 10 items related to program I was doubled and used for the topic score. For "Emphysema" the S's score on the 10 items related to program II was doubled and used for the topic score. To make the data comparable, these two adjustments were made for all three cities. Examination suggested these adjustments were the most conservative which could be made, because they reduced the slight differences indicated by the unadjusted scores.

Two analyses were made of the learning data. The initial analysis involved 18 means, 9 for Ss viewing as a group in a hospital, and 9 for Ss who viewed individually at home. To conserve space, group means will be presented later, and the analysis of variance data will not be presented in tabular form.

In the initial analysis, the F-ratio for (A), Presentation-Response conditions, was 2.10. Although it approached $F_{.90} = 2.30$ at 2/206 d.f., this ratio was not considered statistically significant. Numerically, the means of the overt, covert, and conventional conditions ranked in the order predicted by the hypothesis.

Main effect (E), Conditions of Viewing, had a highly significant F-ratio of 7.49, exceeding $F_{.99} = 6.64$ at 1/206 d.f. The mean was greater for Ss who viewed individually at home, as predicted by hypothesis I.

Statistically significant F-ratios were found for main effects (B), Order, and (C), Topic, both of which exceeded $F_{.999} = 6.91$ at 2/206 d.f. Also statistically significant was the interaction D x E. These findings were duplicated in the final analysis and will be described more fully below.

The major findings of the initial analysis was the support of hypothesis I that learning is greater for general hospital staff nurses who view a continuing education broadcast individually at home than for those who view the same broadcast as a group in a hospital. In contrast, the difference in learning gained from active student participation had little if any practical value for the Ss as a whole. Statistically, the findings on the presentation-response conditions were inconclusive. The program reversal had depressed the overt and covert scores for Ss in City B. These factors, as well as the results of the pilot study, suggested that there might be significant differences for those Ss who completed all overt responses.

To check this possibility, a new variable was added in the final analysis: (F) Response Completion, i.e., response complete and response not complete. Response complete included all Ss who completely fulfilled the specifications for the overt condition for both broadcast periods. Response not complete included all other Ss. The categories response complete and not complete accounted for 31 and 24 Ss who viewed individually at home and 36 and 18 Ss who viewed as a group at a hospital. A Chi-square test showed no significant difference between these groups, indicating that support of hypothesis I was not an artifact of unequal numbers of Ss in these new categories. The 36 means in Table 4 resulted from the dividion into the two new categories of the 18 means previously employed. There were large differences between conditions of viewing, but the differences between presentation-response conditions appeared small.

Table 4

Mean Learning Achievement Scores by
Conditions of Viewing for Three Presentation-
Response Conditions

	Individual Viewing at Home			Group Viewing at Hospital		
	Response Completed	Response not Completed	Both	Response Completed	Response not Completed	Both
Number of Subjects						
(City A)	(16)	(3)	(19)	(14)	(4)	(18)
(City B)	(5)	(15)	(20)	(13)	(7)	(20)
(City C)	(10)	(6)	(16)	(9)	(7)	(16)
(Total)	(31)	(24)	(55)	(36)	(18)	(54)
Overt Condition						
City A	11.25	11.33	22.58	10.57	7.00	17.54
City B	8.40	8.13	16.53	10.77	6.86	17.63
City C	14.20	12.00	26.20	11.56	8.00	19.56
(Total)	33.85	31.46	65.31	32.90	21.86	54.76
Covert Condition						
City A	9.63	10.33	19.96	10.29	8.75	19.04
City B	10.80	8.13	18.93	10.15	7.71	17.86
City C	12.40	10.00	22.40	10.22	8.86	19.08
(Total)	32.83	28.46	61.29	30.66	25.32	55.98
Conventional Condition						
City A	8.63	12.00	20.63	9.00	8.00	17.00
City B	11.40	10.47	21.87	12.23	10.4[illegible]	22.66
City C	11.60	8.00	19.60	8.22	5.71	13.93
(Total)	31.63	30.47	62.10	29.45	24.14	53.59
Grand Total	98.31	90.39	188.70	93.01	71.32	164.33

The final analysis began with the arrangement of the means into four 3 x 3 Graeco-Latin squares, one for each condition of viewing and one for each category of response completion, following a procedure described by Scheffe, who also described the procedure for obtaining the adjusted error term (Scheffe, 1959, pp. 269, 362).

Table 5

Analysis of Variance in Learning Achievement Scores

Source of Variation	Sum Sq.	d.f.	Mean Sq.	F	Significance
(A) Presentation-Response conditions	.82	2	.41	.39	
(B) Order	16.58	2	8.29	7.90	c
(C) Topic	17.19	2	8.60	8.19	c
(D) City	1.27	2	.64	.61	
(E) Condition of Viewing	16.50	1	16.50	15.71	c
(F) Response Completion	24.36	1	24.36	23.20	c
A x E	1.16	2	.58	.55	
B x E	1.47	2	.74	.70	
C x E	1.68	2	.84	.80	
D x E	11.53	2	5.77	5.50	b
E x F	5.27	1	5.27	5.02	a
A x F	2.02	2	1.01	.96	
B x F	.29	2	.15	.14	
C x F	3.52	2	1.76	1.68	
D x F	8.36	2	4.18	3.98	a
A x E x F	2.48	2	1.24	1.18	
B x E x F	.18	2	.09	.09	
C x E x F	.90	2	.45	.43	
D x E x F	5.08	2	2.54	2.42	
Total	120.66	35			
Error (adj.)		194	1.05		

[a] Exceeds F.95 = 2.99 at 2/194 d.f.; 3.84 at 1/194 d.f.

[b] Exceeds F.99 = 4.60 at 2/194 d.f.

[c] Exceeds F.999 = 6.91 at 2/194 d.f.; 10.83 at 1/194 d.f.

For the main effect (A), Presentation-Response conditions, the slight tendency toward significance in the initial analysis completely disappeared in the final analysis. Since the differences were so negligible, this finding would not have changed had there been no program reversal in City B. The statistically significant difference which was found in the pilot study was not found in the main study.

A direct check on the possibility that differences under overt, covert, and conventional conditions would be found for Ss in the response complete category was provided by the first order interaction A x F. No differences were observed.

The effect of (B), Order, was of high statistical significance. The score means indicate the Ss' learning achievement was highest for the first pair of broadcasts in the series. Mean scores declined as the series progressed, although regularity of attendance remained the same. There was no interaction either with condition of viewing or with response completion.

The F-ratio for (C), Topic, was also significant. Means were highest for "Cerebral-Vascular Accidents" and lowest for "Emotional Illness". Little practical significance was attributed to these differences, because while they might have been produced by the instruction, they could also have been produced by differences in topic test difficulty and/or differences in the Ss' topic knowledge at the beginning of the series.

No significant differences in learning was observed for the main effect (D), City, and the only significant interaction was D x E. This occurred because scores in City B slightly favored the condition of group viewing in the hospital, unlike the marked superiority of individual viewing in cities A and C.

The main effect (E), Condition of Viewing, had an F-ratio of increased statistical significance in the final analysis, providing even stronger support for hypothesis I.

The greatest differences between means were related to main effect (F), Response Completion. Smaller, but significant F-ratios were obtained for first order interactions E x F and D x F. Ss in the response complete category scored much higher in learning achievement than the other Ss. Because all Ss in this category completed the overt response condition exactly as prescribed, it might be assumed that this active student participation produced the higher scores. Examination of the means for the non-significant first order interaction A x F indicates this was not the case. The Ss in the response complete category also scored higher than the others under the conventional condition, in which no active response was encouraged. The more acceptable interpretation would be that Ss who score higher in learning achievement are more likely to comply fully with any directions pertaining to instruction.

The major findings of the final analysis were rejection of hypothesis II, that learning achievement varies in descending order for the overt, covert, and conventional presentation-response conditions, and acceptance of hypothesis I, that conditions of home viewing are superior to group viewing in the hospital.

Discussion

The finding of superior achievement for the individual who views at home has relevance for all health science professions. If only group viewing at a central location were effective, broadcast television would have a limited instructional potential. The findings of this study indicate that, through home viewing, broadcast television has the potential to serve all health science professions, including those in which members spend their working day separated from each other.

The results of the pilot study are consistent with previous findings on student

participation. However, in the main study the numerical superiority of learning scores for the overt presentation-response condition is reduced to an amount which can be accounted for by chance alone. This finding is not inconsistent with that of Gropper and Lumsdaine (1961). Consideration of the variables that produce this effect should include these differences in methodology: in the main study the learning measured is that retained for an average of three and a half weeks instead of for one half hour after viewing; the feedback is received by the Ss through the mail from three to five days after viewing instead of immediately after the program; and the Ss are highly motivated, paying volunteers instead of Ss paid to participate.

Some elements which were present in the television course which achieved a high rate of attendance are: 1) subject matter designed to serve the needs of a specified group within a profession, 2) a group of students selected for high motivation, enhanced by the requirement of a fee, and 3) the use of reminder information before each broadcast to encourage each individual to view.

Summary

This study is an experimental investigation into how television can best bring continuing education to health sciences professionals located great distances from medical centers.

One major decision in the use of television concerns the conditions of viewing: whether broadcasts are more efficient when viewed by an individual at home or by a group at a hospital. Another problem concerns which presentation-response condition is most efficient: conventional televised instructions, a covert response by the learner to participation questions inserted at intervals into the broadcast, or an overt written response to participation questions which is reinforced by knowledge of results supplied to the learner through the mail before the next weekly broadcast in the series.

A pilot study used group viewing of videotape recordings by staff nurses at a local general hospital to compare the conventional with the overt condition. These findings supported the pilot study hypothesis that learning achievement is superior under the overt condition, as measured by an external examination following the presentation, when the participation questions were related to the examination questions.

The main study tested the two conditions of viewing and the three presentation-response conditions using television stations in each of three widely separated cities. Both regularity of attendance and learning achievement of the Ss were measured. For general hospital staff nurses, both those viewing individually at home and those viewing in groups at hospitals reported an average of 90 per cent attendance per broadcast. Actual attendance, which could be measured only for hospital viewing, averaged 79 per cent per broadcast. Regularity of actual attendance did not differ for subject matter topics, order of presentation, cities, or presentation response conditions. The Es concluded there were no differences in actual attendance between Ss under the two conditions of viewing.

The initial analysis of learning achievement scores supported hypothesis I, that learning is greater for general hospital staff nurses who view a continuing education broadcast individually at home than for those who view the same broadcast as a group in the hospital. Statistically significant differences were also found for the effects of three orders of presentation and three subject matter topics.

A more detailed final analysis, considering the additional variable of response completion (i.e., Ss who fully completed the overt response requirements compared with Ss who failed to complete the requirements), also supported these findings, but it led to the rejection of hypothesis II that the overt, covert and conventional

presentation-response conditions would vary in descending order. Significant first-order interactions were found between conditions of viewing and response completion, conditions of viewing and cities, and response completion and cities. There were no significant second-order interactions.

Implications of these findings for all health sciences professions and for attendance at television continuing education courses were discussed.

III. MEASURING ACHIEVEMENT IN SCIENCE REASONING ABOUT PHYSIOLOGY*

This study premised that improvement in science reasoning is a major instructional objective in basic health science courses. It is as important to measure student gains of this type as it is to measure gains in specific information about the subject matter. This experiment evaluated a new test of science reasoning about physiology and then used the test to compare the televised instructor-demonstration with the laboratory exercise as a means of improving student performance in science reasoning.

When this study was initiated there was no generally accepted science reasoning test related to the health sciences. The best available instrument had a level of difficulty appropriate to students in the baccalaureate degree program. It was prepared by Paul Dressel and associates (A Test of Science Reasoning and Understanding, 1953) on behalf of, and with the support of, the American Council on Education. Dressel defined science reasoning in terms of specific ability on the part of the student to recognize and state problems; select, evaluate and apply information in relation to the problems; recognize, state and test hypotheses; recognize and evaluate conclusions, assumptions and generalizations; and recognize and formulate attitudes and take action after critical consideration (Instructors Manual for the Test of Science Reasoning and Understanding, 1953). This definition was accepted for the present study.

Review of Pertinent Literature

The relationship of science reasoning with course content, instructor ability and technique of instruction has been investigated. Course organization and instructor ability were found to be important, but research could not identify specific factors that seemed operative. In the related area of critical thinking, studies which made minor changes in particular courses resulted in no greater growth in critical thinking in social science than in unchanged conventional courses (Dressel & Mayhew, 1954).

More positive results have been obtained in studies of techniques of instruction. Science reasoning seems to be best taught by laboratory exercise or discussion. In a review of research studies on techniques of teaching, McKeachie (1962) found that the laboratory exercise appears superior to other techniques in producing ability to apply learning as well as skill in observation and manipulation of materials. Barnard (1942) compared the discussion and the instructor-demonstration in a college science course and found discussion superior on measures of problem-solving and scientific attitude.

The instructor-demonstration appears best suited to teaching principles and specific information. On these measures the Barnard study found the instructor-demonstration superior to the discussion (Barnard, 1942). No direct comparison of the instructor-demonstration with the laboratory exercise has been made in terms of science reasoning, but Cunningham (1946) has reviewed 40 studies which used other measures to compare these techniques of instruction. He concluded that the instructor's precise performance of a difficult experiment may (a) reveal an underlying principle with greater clarity than the student's attempt to perform the experiment for himself, particularly when learning is involved, complicated or difficult; or when the apparatus is complicated or difficult to operate, and (b) convey a greater amount of information in a given time (Cunningham, 1946).

The studies reviewed by Cunningham in 1946 predated closed-circuit instructional television; they compared the individual laboratory exercise with an instructor-demonstration before a small group of students. The performances of a group of 16 students who observed a demonstration of the effects of pharmacological agents in person and of a much larger group who observed via television were recently compared by McGuire, Moore, and Harrison (1961). They concluded that television does not detract from

*Norman C. Staub and Irving R. Merrill

efficacy or quality of the demonstration.

The Problem

The recent interest in televised basic health science instruction raises the problem of whether increased use of televised instructor-demonstrations during a course, perhaps at the expense of a conventional laboratory exercise, has an adverse effect on student improvement in science reasoning. An answer requires prior evidence that improvement in science reasoning about particular course content can be measured. Therefore, this investigation of the problem posed the following questions: (a) can a satisfactory test be devised to measure science reasoning ability concerning control of respiration, and (b) do groups which are taught similar material by televised instructor-demonstration or laboratory exercise differ on improvement in science reasoning?

Method of Study

Three groups of physiology students were equated on science reasoning ability, given a brief textbook assignment on the control of respiration, exposed to differing instructional techniques, and administered a four-part test:

Group I read five reports of original experimental studies in the departmental library (Adrian, 1933; Comroe, 1943; Heyman, 1951; Hutt, Horvath, & Spurr, 1958; Keele & Neil, 1961).

Group II performed laboratory exercises which were modifications of the original experiments.

Group III observed a televised demonstration of the modified experiments by the instructor.

Ss were 63 second-year students in the University of California School of Pharmacy enrolled in Physiology 105, a course offered by the Department of Physiology, School of Medicine.

Prior to the experiment the Ss were given a Test of Science Reasoning and Understanding -- Biological Sciences (1953) and systematically assigned to three groups of 21 students each. Mid-term physiology grades, when compared by analysis of variance, were similar for each group. Because all Ss had received instruction via televised demonstration, laboratory exercise and library reading assignment during the first half of the term, each group was experienced in learning through the technique of instruction it was to receive in the experiment. In order to isolate the effects of the techniques, no lecture was given on the control of respiration.

The experiment was performed within regularly scheduled laboratory periods after mid-term, and the test was given five days later. Ss were not permitted to ask questions during the experiment, because there was no insurance that the same questions would be asked in each group. Each group received six hours of instruction. The design of the experiment eliminated any presumed advantage of the televised demonstration which might be expected on a conventional measure of learning. The relative ease of the laboratory exercises eliminated the possible superiority of the demonstration in communicating the principles involved, and the presentation of the demonstration at a pace convenient for laboratory work eliminated the possibility of communicating a greater quantity of information.

A visual test of visual learning can result in higher scores than a verbal test of visual learning, and vice versa, according to Lefkowith (1955). His findings would predict that Group I, which received instruction by print, would score higher than

Group III, which received it on television, solely on the basis that the test more closely resembled the presentation. To account for this possible artifact, the Ss were given an objective test which included four parts and called for a choice among four answers to each item:

Part I--A 24-item Test of Specific Information--Verbal Choice (I-V) covered items of specific information presented in textbook assignment, prepared by one of the Es, and reviewed by an expert from the Department of Physiology faculty. The form of the items strongly resembled those appearing in the physiology examination prepared by the National Board of Medical Examiners.

Part II--A six-item Test of Specific Information--Pictorial Choice (I-P) was based upon figures in the text of kymograph recordings indicating rate and depth of breathing under particular conditions. The small number of items in this part was dictated by the few illustrations in this portion of the text.

Part III--A 12-item Test of Science Reasoning--Verbal Choice (R-V) strongly resembled in format the test prepared by the American Council on Education for the biological sciences.

Part IV--An 18-item Test of Science Reasoning--Pictorial Choice (R-P) was answered by illustration of curves actually developed during demonstration and in the laboratory.

Thus, the instrument consisted of 30 items on specific information (I-V + I-P) and 30 items on science reasoning (R-V + R-P) and could be combined into 36 verbal choice (I-V + R-V) and 24 pictorial choice (I-P + R-P) items.

Results

Evaluation of Test

All parts of the test satisfied the customary criteria for level of difficulty -- the average S answered about half the questions within each part correctly, and no S either missed all items or answered all correctly within any part of the test.

The reliability of the science reasoning items (R-V + R-P) was considered adequate, although this combination was too brief in its present form for general use. Based upon a measure of internal consistency developed by Hoyt (1941), the coefficient of reliability for the 30 items was .62, whereas for a standardized test of the National Board of Medical Examiners a coefficient of .90 is considered desirable. If the assumptions behind the Spearman-Brown Prophecy Formula (Walker & Lev, 1953) were met, an increase to 54 items would produce the desired coefficient. A 54-item test of science reasoning appears quite acceptable: the test for the biological sciences contains this number and can be administered in less than one hour.

The reliability coefficient of .42 for the combination of specific information items (I-V + I-P) indicated the need for caution when group means were compared on this measure. The combinations of verbal choice items (I-V + R-V) and pictorial choice items (I-P + R-P) produced reliability coefficients of .55 and .54, respectively.

The validity of the 30 specific information items was accepted after two departmental faculty members agreed that the questions were appropriate to the course objectives, the choices reflected specific points at which students frequently go astray, and the answer selected by the instructor was correct. This procedure was also applied to the 30 science reasoning items.

Further evidence of validity was possible from examination of the relationships

among three of each S's scores: science reasoning (R-V + R-P), specific information (I-V + I-P) and the standardized science reasoning test for the biological sciences. The correlation between science reasoning and the biological sciences test should be moderately high and positive, indicating both the common element of science reasoning and the presence of distinctive material which may pertain directly to physiology. If the coefficient were to approach 1.0, all items would be assumed to apply to science reasoning about biology in general. If the coefficient were 0 or negative, the new test (R-V + R-P) would not appear related to science reasoning about any subject. The correlation was r = .57, which was interpreted as moderately high.

The correlation between specific information and the standardized biological sciences test should be low and positive -- the low value indicating major differences between the two measures, and the positive sign reflecting that ability to memorize facts and ability to do science reasoning are both related to general intelligence. The correlation was positive at r = .27, which was interpreted as low. Both correlations were statistically significant at the .05 level.

The predictions that one correlation should be moderately high and that the other should be low implies that the two relationships are different. Analysis according to a formula developed by Hotelling (Walker & Lev, 1953) produced a value of "t" = 2.62, which indicated the difference between the two correlations was statistically significant at the .05 level.

Because the 30 science reasoning items (R-V + R-P) reach at least the minimum criteria for level of difficulty, reliability and validity, it was found that a satisfactory test can be devised to measure science reasoning ability concerning control of respiration.

The Televised Instructor-Demonstration

There was close agreement of the means on all parts of the test (Table 1) for the groups taught by laboratory exercise or televised instructor-demonstration. Means for the control group taught by library reading assignment were lower in seven of eight possible comparisons.

Table 1

Mean Scores by Experimental Groups on 4-Part Test

	Number of Items	Group I Reading	Group II Laboratory	Group III Television
Specific Information -	30	13.48	15.52	14.67
I. Verbal (I-V)	(24)	(11.05)	(12.14)	(11.48)
II. Pictorial (I-P)	(6)	(2.43)	(3.38)	(3.19)
Science Reasoning -	30	14.14	15.09	15.86
III. Verbal (R-V)	(12)	(5.52)	(6.71)	(6.86)
IV. Pictorial (R-P)	(18)	(8.62)	(8.38)	(9.00)
Total	60	27.62	30.61	30.53

Direct analysis of the mean scores for each group on each of the four parts of the test was neither possible nor desired. Of importance was the relationship between techniques of instruction, objectives of instruction, and types of test. Therefore, each S was given four T-scores which were then analyzed by the statistical procedure known as multivariate analysis (Kendall, 1961).

where T_1 = I-V + I-P + R-V + R-P (sum of scores on the four parts);

T_2 = (R-V + R-P) - (I-V + I-P);

T_3 = (I-V + R-V) - (I-P + R-P); and

T_4 = (I-V + R-P) - (R-V + I-P).

The variance and covariance of the several T-scores were first considered. If these were equal in all cases, $-2 \log_e$ lambda would not exceed the value of the Chi-square statistic. The results were as follows:

$-2 \log_e$ lambda = 5.5331 $\qquad$ $\chi^2_{(20\ d.f.)\ .05} = 31.410$

Equality was demonstrated.

Differences between all possible means were then compared. The *F*-ratio served in this case as the test statistic. The results were as follows:

lambda = .64552 $\qquad$ $F_{(8,114)\ .05} = 2.02$

Because the value of *F* was not exceeded, the differences between any of the means of the T-scores were accounted for by chance alone. The science reasoning scores did not differ according to technique of instruction nor did they differ according to type of objective question, i.e., verbal or pictorial choice. The same findings were true for specific information scores.

A careful examination of individual test items was prompted by the finding of no difference which favored the laboratory technique. Two questions might be asked of each item in the test: (a) does it discriminate between Ss of higher and lower ability, and (b) does it discriminate between groups which receive training by different techniques of instruction? If the item failed to do the former, it could not be relied upon to do the latter. The indicator developed by Lawshe (1948) was used to answer the former question, and the Chi-square test was used to answer the latter question.

As anticipated, items which were not sensitive to differences in S ability did not reveal differences related to technique of instruction. Of the 19 specific information and 23 science reasoning items which did discriminate ability, only one showed differences relating to technique of instruction. This was a science reasoning item, and the group taught by television scored higher on it than did the other groups. That none of the other items which were sensitive to ability differences favored any particular instructional technique supported the finding of the multivariate analysis that the laboratory exercise was not superior to the televised instructor-demonstration in improving science reasoning under the conditions of this experiment.

Discussion

There are several applications for the finding that it is possible to prepare a satisfactory test of science reasoning about physiology which a student can complete in less than one hour. An obvious step is the development of equated forms for use at the beginning and end of a course to provide a measure of student improvement in science reasoning during the term. A test having equated forms would permit more

extended experimentation in the teaching of science reasoning. Identification of the relevant variables is worth the continued attention of faculty members in all health sciences. Experience in the social sciences indicates that efforts to teach science reasoning in one course will bear little fruit if other courses in the same curriculum do not share this objective.

Three inferences may be drawn from the finding that the televised demonstration was as effective as the laboratory exercise in improving science reasoning during one unit of a physiology course:

(a) It does not support the assertion that the laboratory exercise is always superior in teaching science reasoning.

(b) It suggests that techniques of instruction are not always a determinant of improvement in science reasoning. A corollary argument is that a certain latitude may be permitted in the substitution of one technique for another during some units of the course. Even the control technique of the library assignment was not less effective than the experimental techniques.

(c) It does not imply that all laboratory exercises may be eliminated from a basic science course without impairment of science reasoning.

Groups II and III did not score higher than Group I on pictorial items, as had been predicted. This finding does not necessarily conflict with the previous studies, which had been done with tests of specific information. The lack of reliability of the six-item pictorial choice test of this type casts doubt on this finding. Other factors may outweigh the importance of presentation-test similarity in the measurement of science reasoning.

Summary

This experiment evaluated a test of science reasoning about physiology and then used the test to compare the televised instructor-demonstration with the laboratory exercise as a means of improving student performance in science reasoning. Three groups of physiology students were equated on an existing test of science reasoning ability about the biological sciences, exposed to differing instructional techniques, and given a four-part objective test.

The results indicated that it is possible to construct a test of science reasoning about physiology which meets the minimum criteria for reliability, validity and level of difficulty. The potential of the test for improvement of science reasoning was discussed.

The televised instructor-demonstration appeared as efficient as the laboratory exercise in improving science reasoning about the control of respiration. The control group of 8s, who read articles reporting the original research on which the instructor-demonstration and laboratory exercises were based, had a lower mean score than other groups, but the difference was not statistically significant.

The hypothesis that the demonstration and laboratory groups would score higher than the control group on parts of the objective test which consisted of pictorial choices was not confirmed.

IV. EFFECTIVENESS OF TELEVISION WITHIN THE DENTAL LABORATORY*

Do students forget part of what they learn from a lecture if they must wait several days before they can apply it in the laboratory? The faculty of the Crown and Bridge Section of the Division of Prosthodontics at the University of California suspects that they do, certainly in the instance when the lecture consists of instructions on a full crown preparation. Forgetting is the familiar word for what is known technically as loss of retention. If S is asked to perform a task he learned previously to a specified criterion of achievement, and if he now accomplishes only 70 per cent of the task, he is said to have 70 per cent retention -- forgetting has occurred. It might be argued that this situation is only a theoretical possibility, for in actual practice students have the opportunity in the interim to refer to books, laboratory manuals, and models, and to consult with advanced students. What they forget may be compensated for by these additional learning experiences.

The major problem of this study was to determine if, under these conditions, forgetting occurs to any measurable extent. If so, the ideal solution would be the elimination of any delay between the lecture and the laboratory application: to have the student instructed while he is at the laboratory table so he can do the procedure immediately. This plan was tried by the faculty several years ago, by moving the lecturer and slide projector into the laboratory with the class of 70 students. The attempt was unsuccessful, due to the shape of the room, the necessity for complete darkness during slide projection, and obstructions to vision by the equipment needed for subsequent laboratory work. The advent of television at the University of California Medical Center at San Francisco made possible an alternative mode of presentation within the laboratory.

The second aspect of this study was to compare student achievement derived from a televised demonstration with that ensuing from the conventional face-to-face illustrated lecture. The technical feasibility of television in the laboratory has been demonstrated for some time (Olson, 1958), but the effectiveness of instruction by television under these stipulated conditions has not been ascertained. The literature reveals only one published experimental study in dentistry on learning through television. This study concerned a postgraduate course in periodontology, and the group taught by television lecture achieved scores as high as the group taught by the conventional method (Tannenbaum, 1956).

In other areas of instruction extensive research on many aspects of this general problem has been done. Some investigators have compared the relative effectiveness of instructional techniques: the lecture with the demonstration. Others have compared similar techniques under two modes of presentation: television with face-to-face. Some have examined visual differences in two presentations: color in photographic slides with monochrome in television, or the static element of slides with relevant motion in television. Other factors that have been investigated are the possible interference to learning from television lights and cameras, and whether students in the viewing room differ in achievement from those in the television originating room. No consistent differences in student achievement have been observed in any of these studies (Eckert, 1960; Carpenter & Greenhill, 1958; Macomber & Siegel, 1960; Jackson, 1952; Rosenstein & Kanner, 1961).

Because none of these factors seemed to be of critical importance, it was decided to analyze the results in terms of the difference between two now generally accepted ways of teaching, the illustrated face-to-face lecture and the television demonstration, bearing in mind that any or all of the factors may contribute to the result.

The purposes of the present study were (1) to determine if a statistically significant

*Theodore S. Grant, Raymond L. Blancheri, Stanley F. Lorencki, and Irving R. Merrill

loss of retention of instruction on crown preparation takes place during a two-day period in sophomore dental education, and (2) to compare achievement between groups receiving instruction on crown preparation by televised demonstration and by face-to-face lecture illustrated with colored slides.

Design and Procedure

The research design in Figure 1 was that of an experiment conducted in a setting of on-going instruction. The mathematical model called for a two-way analysis of variance. Subject matter of the experiment was instruction on the preparation of a tooth in a typodont for a full crown with porcelain veneer facing. The same instructor presented the information to each of the groups of 15 Ss at the point during the semester previously scheduled for this unit. All groups received the instruction at the same time of day, 8:30 a.m. to 8:59 a.m., within a one-week period. Performance tests began in the laboratory at 9:00 a.m. and continued until each S had completed the test case. No S took more than two hours and 45 minutes.

Figure 1

Research Design of Study

Type of Activity	Performance Test Administered	
	Immediately	After Two Days
Televised Demonstration	Group I	Group II
Illustrated Lecture	Group III	Group IV

Group I received its instruction from television monitors located in the laboratory and began the test case immediately after instruction was completed. The camera equipment and the instructor were situated in front of the laboratory. Group II simultaneously received the identical instruction from television monitors situated in the regular classroom, but did not begin the test case for two days. Group III received an illustrated lecture with the instructor, slide projector, and screen physically present in the laboratory, and began the test case immediately. The 15 Ss were able to see and hear satisfactorily. Group IV received the conventional illustrated lecture in the regular classroom, but did not begin the test case until two days later. Both television groups received identical instruction, whereas each lecture group received a separate presentation of the same material.

Despite controls, the experimental procedure was not executed entirely as planned, because only portions of the live televised demonstration approached the standards of technical and instructional quality achieved during the rehearsals. A conference immediately after the demonstration found the instructor, section faculty, television staff, and Es in agreement in this respect. In addition, it was agreed that, although both lectures were of higher quality than the demonstration, the two differed between themselves. Subsequent reference to tape recordings of both experimental lectures showed their differences to be in amount of time devoted to particular topics as well as in other aspects. The consensus at the conference was that the lecture presented to Group IV was superior to that presented to Group III. Analysis of the data verifies this observation.

Subjects. The Ss consisted of sophomore dental students in the course, "Crown and Bridge Prosthetics 423B," at the University of California, School of Dentistry. One student assigned to Group II did not appear for the scheduled instruction, reducing that group to 14, and the total Ss to 59. Prior to the experiment Ss were ranked according to mean score on three test cases performed during the previous semester in "Crown and Bridge Prosthetics 423A," and were assigned to groups systematically.

Measurement. Decline in retention of televised instruction after two days of dental education was measured by subtracting the mean of the scores of Group II from the mean of the scores of Group I. The difference between the means was considered to indicate the actual loss that took place during the two days. Other measures are self-explanatory.

The original design called for measurement of both modes of presentation, in order to supplement studies that compared per cent of retention for periods of 8 months, 1-2 years, and 3 years, respectively (Benschoter & Charlos, 1957; Macomber & Siegel, 1960; Paul & Ogilvie, 1955). For each period subject matter retention was clearly unaffected by mode of presentation. Since differential decline failed to appear after such extended periods, it was not expected to appear for a period as brief as two days. Due to dissimilarities of the presentations before Groups III and IV, measurement of retention for the illustrated lecture was not possible. As indicated, there is no evidence to show that the finding for television is not equally representative for the lecture.

The instrument selected was the test case, a type of performance test. The section faculty devised a scoring procedure which proved highly reliable on similar crown preparations, as measured by a coefficient of concordance (Walker & Lev, 1953). Equally important, the instrument had the attribute of validity. Since the purpose of the instruction was to teach the student to perform a preparation, the most valid test of what he had learned would be his score on an actual preparation.

Controls. The televised demonstration and the illustrated lectures were given by the same E (Lorencki), who also scored all test cases. The 39 cases were placed in random order on a large table, with the name and group number of the S carefully masked. The instructor placed a completed score sheet under each typodont and left the room. A different E then removed the masking material from the typodont and recorded S name and group number on the score sheet.

To insure maximum subject matter similarity between modes of presentation, the conventional lecture was tape recorded, a verbatim copy typed, then analyzed. During a subsequent rehearsal of the televised demonstration, the audio portion was tape recorded, then compared with the lecture version. By this additional rehearsal it was hoped to obtain high quality presentations of both types.

A further problem was in the control of the length of two lecture presentations. The lecture was analyzed by topics, and a warning card system to indicate minutes and seconds remaining for each topic was employed during the experimental presentations. For the sake of uniformity, the same warning system was used during the televised presentation.

Results

Reliability. The first step of the analysis was to estimate the reliability of the performance test and its scoring system. A rough gauge of reliability is obtained if an E's score on the test is consistent with scores on previous test cases. Because Ss were systematically assigned to matched groups, each group included 5 Ss from the upper third, 5 from the middle third, and 5 from the lower third of the total of 60 Ss based upon mean score on three previous test cases. (The S missing from Group II was in the middle third.) If the test was reliable, one could expect that when the two modes of presentation were combined, the present scores of the original upper third would be significantly higher than those of the original lower third. In addition, if reliable, the scores of this group would not be subject to interaction between S ranking and time of test. Mean scores for the 20 Ss in the upper and lower third were 91 and 84, respectively. Division of each third according to whether tested immediately or after two days resulted in the two-way analysis of variance in Table 1. The F-ratio for immediate

versus delayed test may be disregarded, since it includes scores affected by the dissimilarity of the two lecture presentations. The F-ratio obtained between upper and lower thirds would occur by chance less than one in every 100 times. No evidence of anything other than chance interaction was found. It was concluded that the test distinguished consistently between students of higher and lower ability and, therefore, was sufficiently reliable to warrant proceeding with the analysis.

Table 1

Variance in Test Case Scores after Instruction on Preparation for Full Crown with Porcelain Veneer Facing

Explanation	Sum Sq.	d.f.	Mean Sq.	F-ratio	$F_{.99}$
Immediate versus delayed test	76	1	76	1.52	$F_{(1,36)} = 7.39$
Upper versus lower third	570	1	570	11.40	$F_{(1,36)} = 7.39$
Interaction	55	1	55	1.10	$F_{(1,36)} = 7.39$
Individuals within groups	1,810	36	50		
Total	2,511	39			

Retention. Groups I and II were used to test retention of the instruction after two days of other dental education. Difference in means between the two groups could not be attributed to variation of instruction, since each group had observed the identical television demonstration. Table II indicates that Group II forgot so much of the instruction on the preparation that its mean score of 81 represented a decline of 8 points when compared with Group I. A decline of this magnitude could have occurred by chance less than 5 times in a 100; thus a statistically significant decline in retention did occur.

Table 2

Mean Scores and Difference in Achievement by Test Time Administered

	Performance Test Administered		Difference	t-ratio
	Immediately Gp I (n = 15)	After Two Days Gp II (n = 14)		
Televised Demonstration	89.07	81.36	7.71	t=2.15 (sig.)

See Walker & Lev, 1953, p. 158

Mode of Presentation. Table III gives the mean scores for all Ss who received instruction by television and by lecture. It includes those who were tested immediately and after two days for each classification. Thus, the entire presentation score for television is based upon the single television presentation, while the lecture classification is based on two separate lecture presentations. A similar situation applies for the mean scores for that portion of the presentation dealing with the labial cut. They are presented because this portion was considered the most satisfactory at the pre-analysis conference by instructor, section faculty, television staff, and Es. The null hypothesis

of no difference in instructional effectiveness of the illustrated lecture and the televised demonstration was not disproved; the difference between the means was so small as to be accounted for by chance alone.

Table 3

Mean Scores and Difference in Achievement by Mode of Presentation

	Illustrated Lecture Gps. III, IV (n = 30)	Television Demonstration Gps. I, II (n = 29)	Difference	t-ratio
Complete Instruction	88.90	85.34	3.56	t = 1.68 (n.s.)
Labial Cut Only $t_{(.975)} = 2.04$	7.93	7.59	.34	t = 0.60 (n.s.)

See Walker & Lev, 1953, p. 157

Discussion

Perhaps the finding of greatest academic importance is the statistically significant amount of forgetting by dental students who wait several days before they apply instruction obtained during a lecture period in the laboratory. When one considers how often this situation happens to a student within a single course and then multiplies it by several courses for the same student, the interference to successful learning by this single factor assumes serious proportions.

Also of importance is the finding that students who viewed an only partially satisfactory television presentation did not differ in achievement from those students who viewed conventional lectures that were more carefully rehearsed and timed in this instance than would be normally feasible. These two findings provide an answer to the teaching problem of the section faculty: it is more effective to televise instructions that can be applied at once then to give a lecture that cannot be applied for two days.

This conclusion raises several interesting possibilities. It may well be that instructional time can be saved, since a brief televised demonstration within the laboratory period shortens actual working time for the student. If the instructions can be eliminated from the lecture period, the lecturer will have additional time for other important material. In some instances, it may be more desirable to have students achieve higher standards of performance within the existing time schedule. These possibilities must be confirmed by test, and are worthy of investigation.

The advantage of closed-circuit television observed in this study was not related to demonstration versus lecture, monochrome versus color, motion versus a static picture, or possible distractions of television cameras and lights. It was related to its capacity to transport or deliver instruction to a location inaccessible to face-to-face presentation -- in this situation to the student at his laboratory table. When used in this manner, television can contribute to more effective instruction in dentistry.

Summary

If dental students are assumed to forget a significant amount of what they learn from a lecture if they wait for two days before applying it in the laboratory, television can aid retention of knowledge by giving a demonstration to the students while they are at the laboratory table so they can utilize the knowledge immediately. The result should be improved laboratory performance, provided the televised demonstration was

delivered as effectively as the conventional illustrated lecture.

Sixty sophomore dental students in "Crown and Bridge Prosthetics 423B" were divided into four matched groups of 15 Ss each to compare learning from a televised demonstration in the laboratory with three other conditions of presentation: television viewing in a classroom followed by delayed application; conventional lecture followed by delayed application; and conventional lecture within the laboratory followed by immediate application.

The performance test of preparation of a full crown with porcelain veneer facing proved to be a sensitive and reliable instrument due to the system of scoring each test case.

A statistically significant decrease in retention of instructions on crown preparation did take place during two days of sophomore dental education.

When achievement was compared between groups receiving instructions on crown preparation by televised demonstration and by lecture illustrated with colored slides, it could not be shown to differ, despite a televised demonstration that was satisfactory for only some of the steps in the preparation.

Students who receive satisfactory instruction by television that can be applied immediately will score higher in achievement than those who receive satisfactory instruction by classroom illustrated lecture that cannot be applied for several days.

V. THE STEP PRESENTATION OF DENTAL TECHNIC INSTRUCTION*

In a recent experiment instruction in crown preparation was presented by television as a 30-minute instructor-demonstration to one group of students in a classroom and to an equated group in the dental laboratory. The latter group, which immediately applied what had been learned, proved significantly more effective in the preparation of a test case (Grant, Blancheri, Lorencki & Merrill, 1962). Previously, instructions in crown preparation had been given in the classroom as an illustrated lecture covering all the steps of the preparation beginning with the mesial cut. The lecture took approximately 30 minutes.

The present study was designed to test similar groups of students as to whether a presentation of a smaller amount of subject matter at one time by means of televised instruction would result in improvement.

Review of Pertinent Literature

There is considerable reported evidence suggesting that a half-hour television presentation is well beyond optimum step size for efficient learning of this type. As early as 1917, studies were made on the relation between the amount of subject matter presented as a unit and the time required to learn it. Since then, many investigations have shown a disproportionate increase in difficulty of learning with increasing amounts of subject matter. McGeoch and Irion (1952) have reviewed a study by McGeoch and Oberschelp on Peterson Rational Learning Problems, and studies by Cook in 1937 in mental addition and subtraction and in a disc transfer problem. For all three types of problems, as the amount of subject matter was doubled, the difficulty of learning was squared or it was cubed.

This relation recently has received renewed attention. The efficiency of teaching machines and programmed texts has been attributed, in part, to presentation of small amounts of subject matter. As yet there is no rule of thumb to indicate how small these amounts or steps should be. Skinner, one of the leaders in research in programmed instruction, insists that the optimal step size is an experimental question (Green, 1962).

The experiment presented in this paper was conducted to determine if a single cut more nearly approaches optimal step size for presentation at one time than a complete preparation. If so, the total time to complete a test case would be less for the student who received instruction in nine separate steps than for one who received a single presentation covering the complete preparation. The experiment had a second purpose; i.e., to determine whether this saving of time to complete a test case is similar for students of higher and lower class standing.

Method

The experimental subjects were 70 second-semester freshman dental students who had received no prior instruction in crown and bridge. The two groups were equated according to the mean score of four dental morphology test cases. Of all the skills taught during the first semester, these were considered most closely related to those required for crown preparation.

These equated groups of students in adjacent dental laboratories were shown the same recorded televised presentation on crown preparation. The Inclusive Presentation Group viewed the entire recording without interruption and then prepared an ivorine tooth as a test case. The Step Presentation Group received instruction for

*Raymond L. Blancheri and Irving R. Merrill

step 1, the mesial cut, and then performed this step. As soon as the slowest student had finished, instructions for step 2 were presented. Instruction continued with the alternation between presentation and application until the test case was completed.

The subject matter to be learned was well adapted for this experimental study, because all steps in crown preparation could be graded after the test case was completed. For purposes of the experiment, the subject matter to be learned was arranged in nine comparatively equal steps of 5 minutes each: mesial cut; distal cut; buccal -- first step; lingual -- first step; occlusal cut; buccal -- second step; lingual -- second step; correction of line angles and chamfer; and finish. Both groups viewed a 16-minute introductory portion of the recording which identified the instruments and described their function. Student performance was observed in terms of time and quality of work. In both groups time to completion for each student for each step was measured by his laboratory instructor. This measure included only the time the student was at work. Elapsed time included all time from the beginning of the televised presentation to the completion of the test case by the last student in the group. The two measures of the time were considered measures of instructional efficiency.

The quality of work score was considered a measure of instructional effectiveness. Each of the nine steps was weighted according to its relative importance to the completed preparation so that the maximum quality of work score was 100. All cases were scored during one session by the instructor who prepared the televised demonstration (Blancheri). Student name and group identification were masked during the scoring.

Results

The difference in mean time to completion between the Step Presentation Group and the Inclusive Presentation Group was highly significant statistically (Table 1), and the saving of 43 minutes of laboratory time by the Step Presentation Group was considered to have important practical significance as well. It was concluded that approximately 5 minutes of instruction on making a single cut more nearly approached optimal step size than presentation at one time of all nine steps of the preparation.

Table 1

Efficiency and Effectiveness of Two Conditions of Televised Presentation of Crown Preparation

	Groups	
	Inclusive Presentation (n = 36)	Step Presentation (n = 34)
Efficiency		
Mean time to completion	121.78 min.	79.18 min.[a]
Standard deviation	23.19 min.	9.02 min.
Total elapsed time	220.00 min.	171.00 min.
Effectiveness		
Mean quality score	47.19	48.53
Standard deviation	19.52	19.19

[a]Difference between means statistically significant at .01 level.

An additional comparison was made in anticipation of the difficulty of applying these experimental results to actual instruction. The time to completion measure included only the time the student was at work; it ignored the time most Step Presentation Group students spent in waiting for their slowest member to complete the step. The elapsed time measure included this waiting period. Instruction with the smaller amount of subject matter was also found more efficient on this measure as the Step Presentation Group used 49 minutes less than the Inclusive Presentation Group.

The mean quality of work score was, numerically, slightly higher for the Step Presentation Group than for the other Group. When tested by "Student's" t, this difference was found to have no statistical significance, and the quality of work of each group was considered equal for both measures of efficiency.

Because time and quality scores had been obtained for each student, it was possible to note whether the student who does the better work also works faster. Separate correlations were prepared for the Step Presentation and the Inclusive Presentation Groups, but neither approached statistical significance. There were about the same numbers of high and low quality scores for each rate of speed. The amount of subject matter presented did not affect this finding.

To determine whether the smaller step was chiefly of benefit to students of low ability, members within each group were ranked according to mean score for four dental morphology test cases. The upper third of both groups was compared on mean time to completion, as was the lower third of both groups. The difference between these two differences was tested according to the formula of Walker and Lev (1953). The smaller amount of material was of equal benefit to dental students of both levels of ability.

Discussion

This experiment has some resemblance to certain studies on whole-part learning. The latter investigations have divided subject matter into parts, noted the time required to learn each part, and then observed the time required to connect the parts into a whole. While the evidence from these studies is conflicting, there is some reason for believing that the whole method is preferable when much time is needed to connect the various parts (McGeoch & Irion, 1952).

In this experiment, the student was not required to memorize the sequence of events, and poor performance on one step did not affect his score on any of the succeeding steps. Emphasis was on the correct performance of each step in the shortest possible time. The findings in this study are considered relevant to determining the optimum amount of subject matter to be learned; they do not apply directly to problems related to whole-part learning.

The finding of no correlation between quality of work scores and time to completion is not unusual for this type of subject matter. Dental instructors have frequently noted that there is little relation between the quality of a student's test case and the amount of time taken to complete it. Since the study was conducted within the setting of on-going instruction, this finding was not unexpected.

This experiment supports the conclusion that crown preparation may be taught more efficiently by television when small steps are presented for immediate application. The highly significant savings in time to completion were possible only through the use of television, because of the capability of transporting an image directly to the laboratory table. Previous efforts to present laboratory lectures to an entire class of dental students had proved impractical, as would a 5-minute lecture step in a classroom punctuated by a visit to the laboratory and a return to the classroom.

Optimum step size for televised presentation of other types of dental instruction will have to be determined experimentally. For extremely complex subject matter, the optimum presentation may be much less than 5 minutes in length. Through television the amount of subject matter presented at one time may be more easily adjusted to the capacity of the students.

Summary

The major purpose of this study was to test whether it is preferable to teach crown preparation by television in one large step or in a number of small steps. A recorded instructor-demonstration was viewed in adjacent laboratories by 70 second-semester freshman dental students divided into the (1) Inclusive Presentation Group, who viewed the entire presentation without interruption and then prepared an ivorine tooth as a test case, and the (2) Step Presentation Group, who received instruction for the mesial cut, performed this step, and when the slowest student was finished, had the distal cut presented. Instruction for the Step Presentation Group alternated between presentation and application until the test case was completed.

Step presentation was found to be more efficient than inclusive presentation, because the mean time to completion of 122 minutes for the Step Presentation Group represented a saving of 43 minutes. Mean time to completion did not include the time spent by the Step Presentation students waiting for the slowest member of the Group to complete the step, so that an additional comparison was made of elapsed time, which was 49 minutes less than that for the Inclusive Presentation Group.

There was no difference between the groups on effectiveness of instruction as measured on quality of work. There was no significant correlation between a student's quality of work and his time to completion; as many good students worked as slowly as rapidly.

The second purpose of the study was to determine whether smaller steps were of more benefit to students of high or low ability. When the improvement in mean time to completion for the upper third of the class was compared with improvement for the lowest third of the class, the step presentation of dental technic instruction was found to be equally efficient for students at both levels.

VI. CAMERA PLACEMENT FOR RECOGNITION OF COMPLEX BEHAVIORS*

This experiment utilized television recordings and a measure of recognition learning to compare the instructional efficiency of the subjective and observer angles of camera placement. Camera placement is a significant variable and should be related to learning of subject matter in health sciences education, especially in nursing, so as to assist large numbers of professional nurses to conduct the daily nursing care conference.

Leading this conference is a complex activity and is usually centered around two major objectives: to keep group members patient-focused in the problem-solving activities which lead to development of a specific plan of care for each patient and to stimulate interest, understanding, and creativity among members of the group and at the same time to aid them in development of skills in group discussion.

The initial step in teaching professional nurses how to conduct a nursing care conference is to provide situations in which the student learns to recognize the complex patterns of interaction and behavior significant to the task of the team leader. A recorded film or television demonstration is frequently presented for this purpose. This paper presents evidence that the camera angle employed in such demonstrations is related to instructional efficiency.

Review of Pertinent Literature

Motion pictures and television have been generally considered more efficient than printed matter in adding realism to vicarious experience, and use of these media has been recommended for the demonstration of perceptual-motor skills in which realism has been important to instruction. If this consideration is correct, then the efficiency of such a presentation is determined in part by the amount of realism it contains.

Roshal prepared two versions of a film on knot tying and presented each version to a different group of Navy trainees (Roshal, 1949). The presentations differed only in the camera angle from which the instruction was filmed. In the subjective presentation the camera replaced the eyes of the instructor as he demonstrated how to tie the knots. In the observer presentation the camera replaced the eyes of the observer across the table from the instructor. The subjective presentation was found more efficient than the observer presentation because more trainees in this group tied the knots correctly. Roshal concluded, "A film will be more effective in teaching a skill if the task is portrayed from the viewing angle of the learner as he will perform the task." This finding strongly supported the realism hypothesis.

For another group of trainees Roshal prepared a third film version which showed two ends of rope forming themselves into a knot with only the strands of rope visible against a neutral background. More trainees in the group viewing this film tied the knots correctly than those of any other group. This finding appeared to contradict the realism hypothesis.

Miller suggested a resolution of this dilemma by introducing the variable of stages of instruction (Miller, 1957). Particularly appropriate to the early stages of instruction was the elimination of all detail except the relevant cues, i.e., pictorial abstraction. This appeared to account for the greatest success of the third version of the Roshal films. Elimination of the view of the hands left the novice sailors free to concentrate solely upon the contours of the knot and its sequence of preparation. Miller added that as instruction proceeded it seemed advisable to gradually change the presentation in the direction of realism. This suggested that if the experiment were repeated

*Betty J. Dietrich and Irving R. Merrill

with trainees who had some proficiency in knot-tying, the subjective presentation group score might have been higher than that of the pictorial abstraction group.

Roshal focused upon skills training and used a performance test to measure learning immediately after instruction. Our study used a measure of recognition learning and asked: (1) What is the relative efficiency of the subjective and observer presentations when the subject matter is behavior which is more complex than perceptual-motor skill? (2) If differences are found, are they retained for at least four weeks? (3) Is the efficiency of either presentation related to prior subject matter training?

Method

The experimental design called for eight groups classified according to (A) Camera Angle, i.e., subjective or observer presentation viewed, (B) Time of Test, i.e., measurement of immediate or retained learning, and (C) Prior Training, i.e., some or no previous formal instruction of Ss in the team concept of nursing care. All Ss were assembled for the experimental procedure, and the group was divided to view the subjective or observer instructional presentation. Half of those who viewed each presentation remained and were given an achievement test to compare immediate learning. The other half was excused and reassembled four weeks later and given the same test to compare retention. Each division of the total group identified by camera angle and time of test contained roughly equal numbers of Ss with some or no prior training.

The subjective and observer versions of the instructional presentation were both kinescope recordings of 22 minutes duration, and were identical except for camera angle during the demonstration portions. Their format was that of a brief lecture in five parts, each part immediately followed by a prepared demonstration. The demonstrations accounted for 15 minutes of each presentation. The same experienced demonstration team participated and the same words were spoken in both versions. Only one camera was used during either presentation. The team leader role was the focus of both presentations, because the instructional objective was recognition of what the effective team leader does to conduct a nursing care conference.

The subjective presentation portrayed the leadership task from the viewing angle of the learner as she (or he) performs the task. The camera unit occupied the team leader's position in the group and the camera turned -- as the team leader would have done -- toward the licensed vocational nurse, nurse aide, orderly, or student nurse as appropriate to reveal pertinent spoken or non-verbal communication. The focus of the camera was fixed to show only the head and chest of one person at a time. The script was prepared and the setting was designed not to distract from the viewer's sense of participation as the team leader.

The observer presentation portrayed the leadership task from the viewing angle of the student nurse, shown in Figure 1 seated opposite the team leader. The camera was focused on the head and chest of the leader throughout the demonstrations, and did not show any other team member. Videotape recordings of both instructional presentations are on file at the Office of Television Research, University of California Medical Center, San Francisco 22.

Measurement of learning retained after four weeks was included in the design to determine whether any differences in efficiency of the two presentations were of more extended duration. The variable of prior training was included to test the hypothesis that realism of presentation is more efficient at the later stages of instruction. Support of the hypothesis requires that students who have had some prior instruction score higher from viewing the subjective presentation than those who have had no prior instruction.

The Ss were registered nurses, employed in staff, head nurse, or supervisory positions, who volunteered to participate. Several weeks before the experiment, prospective volunteers were invited to complete a questionnaire which included questions about formal instruction and practical experience in team nursing. Formal instruction was considered by the Es to be more relevant than practical experience to the type of recognition learning measured in this experiment. Therefore, those Ss included in the some prior training classification had previously received formal instruction in team nursing in a workshop, a continuing education short course, and/or a one-semester academic course. Most of these Ss also had practical experience in team nursing. Those included in the no prior training classification had received no previous formal instruction and few of them had any practical experience in team nursing.

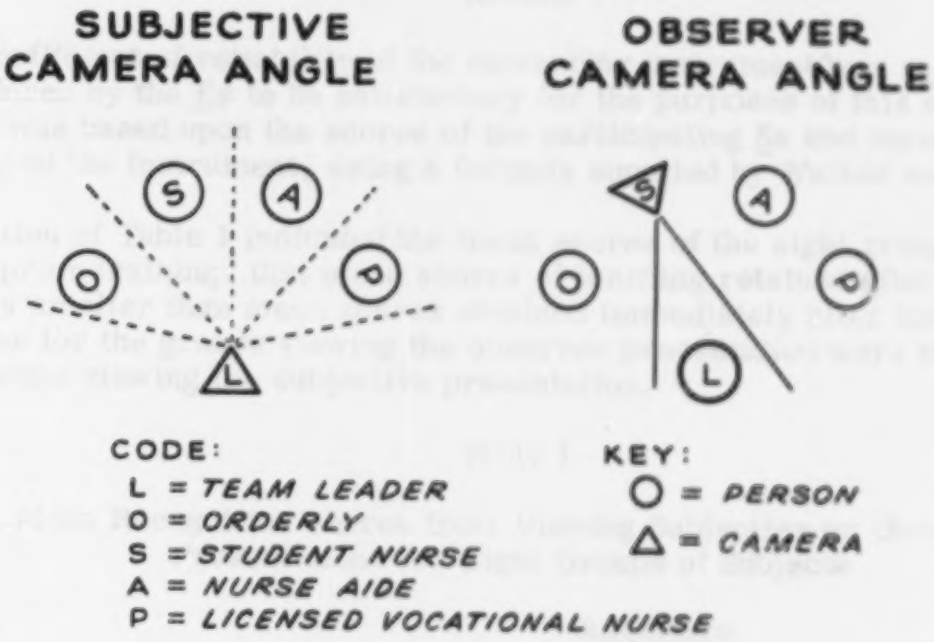

FIGURE 1. SCHEMATIC ILLUSTRATION OF THE SUBJECTIVE AND OBSERVER CAMERA ANGLE PLACEMENT FOR DEMONSTRATION OF THE ROLE OF THE TEAM LEADER IN THE NURSING CARE CONFERENCE.

The number of prospective volunteers who satisfied the some and no training criteria were 33 and 30, respectively. A table of random numbers was used to assign the Ss in each classification to groups according to camera angle and time of test. Chance determined whether the maximum number of potential Ss per group was seven or eight. Fifty-three Ss appeared at the first experimental session, but three failed to return after four weeks for measurement of learning retained, reducing the total to 50. One of those examined for retained learning was excused because she was a recent immigrant who had extreme difficulty in reading the examination questions. Therefore, the appropriate column heads in Table 1 indicate 49 Ss were included in the analysis of results.

The dependent variable in this experiment was recognition learning, measured either immediately or four weeks after instruction. The measuring instrument had a special claim to validity, because it approximated the stimuli of the instructional presentation more closely than could a verbal description within the conventional pencil and paper test (Hartman, 1961). The instrument consisted of a test presentation

demonstrating a typical nursing care conference with an associated multiple choice question sheet. The test presentation was a 28 minute kinescope recording produced in a television studio using conventional two-camera production techniques. The test presentation used a different cast from that which appeared in the instructional presentations. This cast demonstrated a set of actions, some of which reflected and some of which violated the leadership approaches taught in both instructional presentations.

The associated question sheet measured learning by the recognition method because it required the S to indicate whether actions demonstrated in the test presentation were taught in the instructional presentation (McGeoch & Irion, 1952). The question sheet was refined from two previous pilot versions.

Results

The coefficient of reliability of the measuring instrument was $r_{XX} = .775$, which was considered by the Es to be satisfactory for the purposes of this experiment. The coefficient was based upon the scores of the participating Ss and measured the internal consistency of the instrument, using a formula supplied by Walker and Lev (1953, p. 311).

Inspection of Table 1 indicated the mean scores of the eight groups did not vary according to prior training, that mean scores of learning retained after four weeks were numerically smaller than mean scores obtained immediately after instruction, and that mean scores for the groups viewing the observer presentation were numerically larger than for groups viewing the subjective presentation.

Table 1

Mean Recognition Scores from Viewing Subjective or Observer Presentations for Eight Groups of Subjects

	Subjective		Observer	
	Number in Gp.	Mean Score	Number in Gp.	Mean Score
Immediate Learning Measurement				
Some Prior Training	8	8.38	6	14.50
No Prior Training	5	13.20	8	11.25
(sub total)	(13)	(10.23)	(14)	(12.64)
Retained Learning Measurement (after 4 weeks)				
Some Prior Training	7	9.29	7	9.57
No Prior Training	5	7.60	3	13.33
(sub total)	(12)	(8.58)	(10)	(10.70)
Grand Total	25	9.44	24	11.83

The statistical significance of these numerical differences was tested by analysis of the variance. Because there were unequal numbers of Ss in the groups, it was appropriate to employ the test for homogeniety of variances where samples are of unequal size and the approximate method of computation of the sums of squares as described by Walker and Lev (1953, pp. 192, 381). The mean of each group was then treated as a single observation.

The first experimental question involved a comparison of the relative efficiency of the subjective and observer presentations. The main effect A (Camera Angle) in Table 2 indicated that the differences between the mean scores of those groups which viewed the two instructional presentations (11.82 - 9.44) was statistically significant. A difference of this size would be expected by chance less than five times in 100. These results indicated the observer presentation produces more efficient recognition learning than the subjective presentation when the subject matter is more complex than perceptual-motor skills.

Table 2

Analysis of Variance

	Sum of Sq.	d.f.	Mean Sq.	F-Ratio
A (Camera Angle)	12.95	1	12.95	4.69[a]
B (Time of Test)	7.10	1	7.10	2.57
C (Prior Training)	1.65	1	1.65	.60
A x B	.43	1	.43	.16
A x C	.97	1	.97	.32
B x C	.04	1	.04	.01
A x B x C	22.84	1	22.84	8.26[b]
Total	45.86	7	6.55	2.37
Error (adj.)		41	2.76	

[a] Exceeds $F_{.95} = 4.08$ at 1/41 d.f.

[b] Exceeds $F_{.99} = 7.31$ at 1/41 d.f.

The second question asked if this difference in learning persisted as long as four weeks. Main effect A indicated the observer presentation had superior efficiency even when groups which had both times of test and classifications of prior training were included. Main effect B (Time of Test) indicated there was no statistically significant loss of learning among all groups after four weeks. The numerical decrease in mean score noted during inspection of Table 1 was so small as to be attributed to chance alone. However, the possibility remained that passage of time might significantly affect the degree of the observer presentation superiority, because the difference between presentation mean scores for those tested immediately was 2.41 and for those tested after four weeks was only 2.12. This possible relation between camera angle and time of test was considered in the first order interaction A x B. The F-ratio signified that this decrease in the extent of the differences was attributable to chance alone. These results indicated the observer presentation is more efficient than the subjective presentation in producing learning when measured immediately after instruction and that this degree of superiority in efficiency is retained for as long as four weeks.

The third question was whether the efficiency of either presentation was related to prior training. Main effect C (Prior Training) indicated that there was no statistically significant difference between the mean scores of all groups with some and with no prior training. However, the possibility remained that there may have been some interaction between one or both presentations and the prior training variable. First order interaction A x C analyzed this possibility, but the F-ratio of .32 indicated the minor differences observed did not even approach statistical significance. These results indicated that prior subject matter training is not related to the efficiency of either the subjective or the observer presentation. This finding did not support the

hypothesis that the realism in the subjective presentation is more efficient during the later stages of instruction, i.e., for groups with some prior training.

The second order interaction A x B x C accounted for the remaining degree of freedom among the groups. This F-ratio was of high statistical significance, indicating the need for care in interpretation of main effects or first-order interactions. For example, the group with prior training which received the measure of immediate learning and the group with no prior training which received the measure of retained learning together accounted for most of the superiority of the observer presentation. There appeared to be little real difference related to presentations between the other four groups. However, because of the size of the difference between sub-totals for both immediate and retained learning measurements, the statistically significant superiority of the observer presentation was acceptable as a main effect.

Discussion

Before consideration of the application of these findings to nursing education, two comments are in order: (1) This study does not invalidate that of Roshal, because the experiments differed in the subject matter of instruction and the related measure of learning. Roshal presented substantial evidence that the subjective presentation is more efficient for demonstration of perceptual-motor tasks when measured by a performance test. The evidence was equally clear from this study that the observer presentation is more efficient for demonstration of these more complex behaviors, based upon a measure of recognition learning. The development of an hypothesis which relates these two sets of findings deserves attention.

Second, there was a factor within this experiment which may have accounted for lack of support of the realism hypothesis. The questionnaire sent to prospective volunteers asked about formal instruction related to the general topic of team nursing. Such instruction usually touches briefly on some of the approaches the team leader may use in the successful conduct of the daily nursing care conference. However, it was possible that there was no instruction pertaining to this aspect of the team leader's role. If this were the case, the groups would not differ as to stage of instruction and no differences could be expected.

The finding that the efficiency of a televised demonstration was related to camera placement deserves consideration by nursing educators. The present study indicates the observer presentation is more efficient in nursing education where recognition learning is important, such as interviewing and observing patients to identify specific nursing issues or observing and resolving personnel problems. Many educators may wish to explore the efficiency of the subjective presentation in teaching the many perceptual-motor tasks the modern nurse must be competent to perform, from bathing a patient to use of the tonometer. Because much of nursing instruction is highly visual in nature, the search for fruitful applications should not be limited to the course on team leadership.

Summary

This experiment used television to compare the instructional efficiency of two angles of camera placement in the conduct of a daily nursing care conference by the effective nursing team leader. Scores based upon a measure of recognition learning were analyzed in terms of the subjective or observer camera angle, immediate or retained learning, and some or no prior training. Both instructional presentations were identical except for camera placement.

The 49 experimental Ss were registered nurses employed in staff, head nurse, or supervisory positions and were selected from a large group of volunteers. About half

of this number had previously received a limited amount of formal instruction in team nursing; the remaining half reported no previous formal instruction on this topic.

The measuring instrument was designed to closely approximate the conditions of instruction and consisted of a test presentation and associated question sheet. The test presentation was a television recording of a nursing care conference by a demonstration team whose actions sometimes reflected and sometimes violated the leadership approaches taught in the instructional presentations. The associated question sheet then required the S to identify those test presentation actions of the team leader which had been taught in the instructional presentation.

The observer presentation was found to produce more efficient recognition learning than the subjective presentation when the subject matter was more complex than perceptual-motor skills. This superior efficiency was observed immediately after instruction, and the degree of superiority was retained for as long as four weeks. Prior subject matter training was not related to the efficiency of either the subjective or the observer presentations. This finding does not support the hypothesis that the realism in the subjective presentation is more efficient during the later stages of instruction.

APPENDIX A. INSTRUCTOR-TELEVISION STAFF COOPERATION*

Coupled with potential gains from research is the reality of day-to-day operation. Just as the quality of instruction depends in large measure on the instructor, the technical performance of the television system depends upon the quality of the television staff. There is no efficiency in originating with the finest instruction if it is mediated in such a way as to display a blurred image and distorted sound. There are several points of view on the necessity for a staff which is capable both of operation and maintenance. Both parties to this good-natured controversy have the responsibility to insure that their respective students are taught on equipment which is operating at, or very near, its optimum capacity.

Good instruction, of course, is not guaranteed by an able instructor and a competent television staff; teamwork is also required. The health sciences instructor has a major advantage in experience over the instructor in, for example, the humanities. For the humanist, instruction is always an essentially individual performance, involving a sheaf of notes and an occasional piece of chalk. The health sciences instructor organizes his presentation by selecting and arranging a series of photographic slides and film clips, preparing specimens and/or microscopic slides, with the participation of projectionists, laboratory assistants, and assistant instructors. With the addition of television, the health sciences instructor merely expands the team or changes some of its members.

Implicit in this approach is the recognition that for television to dominate the instructor is undesirable. The instructor should at all times be in control of the presentation of his material; however, he will not be able to control what is displayed on the viewing monitors unless he welcomes the technical staff into his confidence. For uncomplicated presentations, he should satisfy himself that staff members understand the objectives of the presentation, and are informed of the sequence in instruction. This permits them to plan the camera shots in advance, selecting the appropriate cameras and lenses. It may be desirable to check out several of the shots on camera to insure that the required visual information is displayed or, on occasion, a complete rehearsal may be feasible. One of these types of coordination is always required, whether the technical staff consists of faculty and student assistants or of full-time television technicians. As this cooperation continues, the instructor will find that the staff will be able to draw upon a background of experience in similar presentations which may help him in selecting the most efficient technique for his own presentation.

Cooperation helps to insure that the student sees what should be seen and hears what should be heard at the time best suited for learning. It is a good rule of thumb to exclude elements which do not directly contribute to this goal.

*Irving R. Merrill

APPENDIX B. PERCEPTUAL RECOGNITION OF PICTORIAL SIGNS*

In this series of four experiments, pictorial signs were presented at several durations of exposure to investigate the effect of familiarity on perceptual recognition. The pictorial signs used in these experiments were line drawings, which were considered to represent certain characteristics of objects in much the same way as the corresponding verbal signs (Osgood, Suci, & Tannenbaum, 1957, pp. 5-9).

Recognition of verbal signs, words either written or spoken, under reduced stimulation has received intensive investigation. In a review of these studies Postman and Rosenzweig pointed out that (a) perceptual recognition is a monotome function of word familiarity; the speed with which a particular word is recognized depends upon the frequency with which this word has been used in the past, (b) both recognition and recall measures of learning reflect the relationship of word familiarity to verbal behavior, and (c) the relationship of word familiarity to recognition is most pronounced in the region of the perceptual threshold (Postman & Rosenzweig, 1957).

Two problems of measurement appear in the studies of word recognition. One pertains to the type of response required of the S. Considering the studies mentioned in the above review, in those which presented visual stimuli, the S was required to give a spoken response. Conversely, written responses were required in the studies presenting auditory stimuli. This raises the possibility that monotonicity of data may have been a function of the response required, not the effect of familiarity on perceptual recognition. Owens reported that under an unfiltered condition of auditory stimuli, errors in writing words of low familiarity are greater than in words of high familiarity at the .01 level of statistical significance (Owens, 1961, p. 122). No familiarity effect was expected during the unfiltered presentation, either upon perceptual recognition or the ability to write the word which was spoken, because all words were three phonemes in length and had been matched as to initial and final consonent sound, e.g., low familiarity -- roan, moderate familiarity -- rain, and high familiarity -- run. Even in this carefully controlled study, errors in writing unfamiliar words appeared where errors in recognition were not predicated. A demonstration that the effect of familiarity on perceptual recognition is separable from the function of the written response would lend additional support to the familiarity hypothesis.

The other problem is one of the similarity of the response to the conditions of presentation and test: the more similar the response, the higher the score. In the two-channel study of Postman and Rosenzweig, written response scores were highest when printed words were used in training and test presentations; written response scores were lower when spoken words were used in training and test presentations (Postman & Rosenzweig, 1956). However, in a recognition study using pictorial signs it is not feasible to require the Ss to draw their responses, because drawings are very difficult to score. An alternative solution is the use of a very simple response which indicates recognition but does not require the Ss to speak, write, or draw. There is the possibility that with a simpler response, briefer exposures may be required to reveal the effect of familiarity. The exact durations of exposure related to this recognition test form must be discovered experimentally.

The purpose of the present investigation was to determine:

(1) Does familiarity affect perceptual recognition of pictorial signs?

(2) At what duration of presentation and test exposure does the familiarity effect appear when the recognition test form is used as a measure?

*Irving R. Merrill and John G. Shrock

(3) Is the effect of familiarity upon perceptual recognition, as measured by the recognition test form, separable from its effect upon ability to write the name of the object pictured?

Method

Design. The design employed in each of the four experiments involved (a) exposure of a sequence of 39 items of a Presentation List, (b) completion of a recognition test during exposure of a sequence of 75 items of a Test List, and (c) completion of an identification task during a second exposure of a Test List. The same lists were used in each experiment. Within each experiment the duration of exposure was identical for both Presentation and Test List.

The major difference between the experiments was the duration of exposure: 3,000 milliseconds for Experiment I; 5 milliseconds for Experiment II; 2.5 milliseconds for Experiment III; and 1.25 milliseconds for Experiment IV.

Postman and Rosenzweig reviewed studies in which perceptual recognition of words was measured after determination of the individual perceptual threshold of each S. In the present investigation, perceptual recognition of line drawings presented at a single duration of exposure for a given group of Ss was measured; individual differences in threshold of perception were not ascertained.

Stimulus Materials. Stimulus materials were 75 line drawings of objects or persons, of which 25 were identified as being of high familiarity, 25 of moderate familiarity, and 25 of low familiarity. Familiarity was defined as the frequency of usage of the corresponding object (noun) as determined by the Thorndike 10,000,000 word count (Thorndike & Lorge, 1952).

Uniform style and picture size of 8" x 5" was maintained for all drawings. In nine pictures of moderate and 10 pictures of low familiarity reference lines were added to avoid ambiguity (Fig. 1).

The Presentation List consisted of 12 drawings of high familiarity, 14 of moderate familiarity, and 13 of low familiarity selected at random from the 25 pictures in each of the respective categories. A table of random numbers was used to assign these 39 items to a sequence of exposure used in all experiments. The Test List consisted of all 75 drawings assigned by a table of random numbers to a sequence of exposure which also was used in all experiments. In Table 1 the object nouns corresponding to the pictures in the Presentation and Test Lists were presented alphabetically for ease of examination.

Subjects. The total number of voluntary experimental Ss was 45. Each S participated in only one experiment. Ss were engineers, professional staff members, nurses, personnel workers, artists, or clerical workers at the University of California Medical Center. Because each experiment was conducted at a different time, and Ss were obtained as needed, it could not be assumed that homogeneity existed among the experimental groups. However, inspection of each group for occupations represented and the proportion of men to women indicated no sharp divergencies among the four groups.

Procedures. Experiment I involved a duration of exposure at which no effect of familiarity could be anticipated. To insure that the duration was above any individual threshold of perceptual recognition an extended duration of 3,000 milliseconds was selected and a group of 22 Ss included. To accommodate this number of Ss the experiment was conducted in a small well-lighted classroom. One of the Es held up each picture as the other checked duration of exposure by stop-watch.

HIGH FAMILIARITY

APPLE

GIRL

MODERATE FAMILIARITY

CHALICE

LAPEL

LOW FAMILIARITY

KIP

MANACLE

Fig. 1. Examples of Pictorial Signs in Each Category of Familiarity.

Note Reference Lines in Illustration of LAPEL and MANACLE.

To determine a duration below any individual threshold, nine Ss were tested individually on a Gerbrand tachistoscope in Experiment II. A duration of 5 milliseconds was selected. Because the smallest marking on the scale for setting the duration adjustment of the tachistoscope was 10 milliseconds, an interpolation setting midway between that mark and the zero-point was required. When all nine Ss reported little difficulty in seeing the pictures, it was decided to conduct additional experiments at durations of 2.5 and 1.25 milliseconds. Nine Ss were also scheduled for each of these experiments. However, two Ss failed to appear for Experiment III; and two Ss in Experiment IV refused to complete the Presentation List, reporting they saw no pictures at durations of 1.25 milliseconds. Thus, the analyses of Experiments III and IV were each based upon observations of seven Ss.

Table 1

Test List by Category of Familiarity
with Presentation List Indicated by Asterisk

High	Moderate	Low
apple	album	abutment
arm*	alcove	ballerina
ball	awning	baluster*
bear*	barricade*	bassoon
bed*	baton	binnacle*
bird*	carcass*	bodkin
book	chalice	brickbat
box	churn*	calipers
boy*	demon	camisole
bread	diadem*	capstone*
church*	dais	caryatid
corn*	emblem	compote
door*	frill*	demijohn*
egg	fuse	fleur-de-lis
eye	girder	gavel*
fly*	glider*	hieroglyph*
foot	harmonica*	joist*
girl	icicle*	juggernaut
hand*	javelin*	kip*
heart	keg*	lasso*
horse	lapel*	manacle*
house*	latch*	parfait*
leaf	lute	piccolo*
letter	mantle*	pyre
man*	niche*	quirt*

Because the latter two experimental durations were much briefer than those employed in previous experiments of perceptual recognition of words, all settings of the tachistoscope were rechecked with an oscilloscope. Settings of 5 and 2.5 milliseconds checked out accurately after repeated exposures. The oscilloscope traces for a setting of 1.25 milliseconds were different from those of 2.5 milliseconds, but were not consistent. Therefore, the setting at 1.25 milliseconds may not be considered accurate, but it can be assumed to be of briefer duration than 2.5 milliseconds.

In the sequence of procedure, each S received a general orientation and instructions on the recognition test. A single pictorial sign, which was not part of any list, was exposed as an example at the experimental duration; then the Presentation List was exposed, followed by the Test List. After explanation of the identification task, the Test

List was exposed for a second time.

For all experiments the interval between exposure of each of the 39 items of the Presentation List was three seconds. A five second interval was maintained between the 75 items of the Test List during the recognition test, and an eight second interval was maintained during reexposure of the Test List to permit the Ss to write the name of the object illustrated.

Measurement. The form of recognition test used in each experiment did not require the S to write the name of the object illustrated. He merely placed an "x" in an appropriate column on the answer sheet to indicate whether a particular item in the Test List "Did Appear" or "Did Not Appear" in the Presentation List as well. Each category of familiarity was scored separately, with each correct response contributing to the score. Thus, the possible range of recognition scores a S could obtain in any category was 0 - 25.

Identification scores were obtained from the second exposure of the Test List. The S was required to write the correct name of each object illustrated. Again, each category of familiarity was scored separately, and a correct response contributed to the score, providing a possible range within any category of 0 - 25.

Results

Inspection of mean recognition scores, Table 2, indicated that at the extended duration of 2,000 milliseconds the means of the three categories appeared similar, at 5 milliseconds the mean of the high familiarity exceeded the moderate and low familiarity categories, and at both 2.5 and 1.25 milliseconds the means assumed the numerical pattern of the monotone function of familiarity. Identification scores, based upon the correctly written name of the object illustrated, reflected a marked monotonic effect of familiarity at all durations of presentation. This effect was so marked that it required no further statistical test.

Table 2

Mean Recognition and Identification Scores for
Pictorial Signs of Differing Familiarity at Four Durations of Exposure

Experiment	Duration (msec.)	Number of (Ss)	Recognition Scores Degree of Familiarity			Identification Scores Degree of Familiarity		
			High	Mod	Low	High	Mod	Low
I	2,000.00	(22)	23.8	23.5	24.0	22.6	8.1	3.8
II	5.00	(9)	19.0	15.9	17.2	19.2	2.8	1.1
III	2.50	(7)	15.3	13.6	10.4	13.1	1.0	.2
IV	1.25	(7)	13.9	12.3	11.0	4.6	-	.1

The numerical differences noted among the recognition scores were tested by the analysis of variance. If the four experiments had been combined into a single design with, e.g., one duration of exposure for the Presentation List for all experimental groups and various durations for the Test List, then a two-way analysis of variance would be appropriate. The four separate experiments in the present study were each suited to a one-way analysis of variance to test whether differences in recognition score means of the familiarity categories had statistical significance. An analysis for comparison of the means of several measures of the same individuals was appropriate (Walker & Lev, 1953, p. 216). F-ratios for Experiments I through IV were 82, 3.12,

20.22, and 19.53, respectively. The ratios for Experiments I and II were not significant at $F_{.95}$ with 2, 42 and 2, 16 degrees of freedom, respectively. Ratios for Experiments III and IV were significant beyond $F_{.99}$, both at 2, 12 degrees of freedom.

The first experimental question asks whether familiarity affects perceptual recognition of pictorial signs. The familiarity hypothesis is tested by one of the two degrees of freedom between the three category means. If analysis of this single component indicates the means have a linear relationship, then the statistical significance of the familiarity effect is established (Cochran & Cox, 1957, p. 62). Examination of the F-ratios in Table 3 resulting from these four analyses indicated that in Experiments III and IV, the linearity of the means was of very high statistical significance. A difference of this amount would be expected by chance far less than one time in 100. Because a linear function falls within the more general definition of a monotone function, these results indicate that familiarity does affect perceptual recognition of pictorial signs.

Table 3

Summary of Four Analyses of Linearity of Variance of Recognition Scores for Pictorial Signs of High, Moderate, and Low Familiarity

Experiment Number	Duration of Exposure (msec.)	F-ratio	$F_{.95}$	$F_{.99}$
I	3,000.00	.34	4.07	7.27 at 1, 42 d.f.
II	5.00	1.77	4.49	8.53 at 1, 16 d.f.
III	2.50	29.89	4.75	9.33 at 1, 12 d.f.
IV	1.25	26.09	4.75	9.33 at 1, 12 d.f.

Because it was possible that the presence of reference lines in some pictures of the moderate and low categories could have accounted for the familiarity effect, the complete analysis was repeated using only 15 pictures in each category which did not contain reference lines. The results were similar to the first analysis; the familiarity effect was not an artifact related to reference lines.

The second question concerns the relation between the form of recognition test employed in this study and the duration of exposure at which the familiarity effect was observed. Table 3 confirms that no difference was observed in Experiment I at the extended duration of exposure of 3,000 milliseconds, as expected. The numerical differences observed in Experiment II at the 5 millisecond duration were not considered statistically significant, because they could be accounted for by chance alone. It was only at the extremely brief durations of 2.5 and 1.25 milliseconds that the familiarity effect appeared. These findings indicate that if the duration of exposure is the same for all items whenever exposed and if perceptual recognition of pictorial signs is measured by the recognition test form, the familiarity effect occurs at the 2.5 and 1.25 millisecond durations of exposure.

The third question concerns the relation between the familiarity effect revealed by use of the recognition test form and by use of the identification task. If the effects are the same, the monotone function should be present or absent in both recognition and identification scores for any given experiment. If this correspondence does not occur in some or all of the experiments, then the effects are separable. Table 2 indicates the presence of the monotone function in both sets of scores for Experiments III and IV. However, in Experiments I and II the effect was present in identification

scores only. Table 3 indicates the effect was absent from the recognition scores of these two experiments. Therefore, the effect of familiarity upon perceptual recognition as measured by the recognition test form is separable from its effect upon ability to write the name of the object pictured.

Discussion

Postman and Rosenzweig concluded from a review of the literature that the same principles of learning apply both to linguistic behavior and to the perceptual recognition of verbal signs (Postman & Rosenzweig, 1957). This study provides support for the expansion of the statement to include perceptual recognition of pictorial signs.

The exposure at which the familiarity effect was observed was very brief. That it should appear in recognition scores at a duration of 2.5 milliseconds or less was not predicted by previous studies requiring a written response from the S. However, the accuracy of the tachistoscopic settings was confirmed by measurement of the durations by means of an oscilloscope. More than one factor may have contributed to this brevity. As suggested above, a test form which required only an "x" instead of a completely written word for a correct response may have accounted in part for this reduction. Further investigation of the test variable along with other possible variables is indicated.

Considered by themselves, Experiments I and II demonstrated that an apparent familiarity effect may be obtained from a written response when a corresponding effect is not obtained from a form of recognition test. This comparison of recognition and identification scores indicated the importance of design and control in investigation of perceptual recognition under reduced stimulation. However, when the recognition scores of all four experiments were considered together, the presence of the familiarity effect only in Experiments III and IV gave additional support to the previous finding that the familiarity of signs does affect their perceptual recognition.

APPENDIX C -- RESEARCH STAFF

Principal Investigator:	1/1/60-9/30/63	Theo. S. Grant, D.M.D.
Associate Investigators:	1/25/61-9/30/63	L. L. Bennett, M.D.
	1/1/60-6/30/60	J. T. Crane, M.D.
	1/1/60-9/30/63	Betty J. Dietrich, R.N., M.S.
Director of TV Research:	1/1/60-9/30/63	I. R. Merrill, Ph.D.*
Statistician:	2/22/61-6/30/63	J. G. Shrock, Ed.D.*
	2/1/63-9/30/63	Ruby Y. Buenz*
Senior Medical Television Engineer:	8/24/60-9/30/63	R. A. Litke
Television Technicians:	9/21/60-9/30/63	R. A. Mort*
	3/29/61-9/30/63	F. J. Knight
	6/19/61-9/30/63	D. L. Adam
	3/1/62-6/15/62	R. M. Dunn
	11/5/62-11/30/62	J. E. Ross*
	12/13/62-9/30/63	Susan C. Clark*
Laboratory Helper:	10/20/60-6/2/61	R. M. Harkey*
Communications Production Assistant:	11/1/61-9/30/63	H. W. Symmes
Project Secretary:	3/1/60-9/21/60	Tana B. Mitchell*
	10/24/60-7/25/63	Lynne D. Dubiner*

*Salary paid from grant funds during all or part of the period indicated.

REFERENCES

Ackoff, R. L. The design of social research. Chicago: University of Chicago, 1953.

Adrian, E. C. Afferent impulses in the vagus and their effect on respiration. J. Physiol. (Lond.), 1933, 79, 332-358.

Allen, W. H. Research on film use: student participation. AV Communication Rev., 1957, 5, 423-452.

Barnard, J. C. The lecture-demonstration v. the problem-solving method of teaching a college science course. Sci. Educ., 1942, 26, 121-132.

Benschoter, R. P., & Charles, D. C. Retention of classroom and television learning. J. appl. Psychol., 1957, 41, 253-256.

Blancheri, R. L., & Merrill, I. R. Television in health sciences education: II. the step presentation of dental technic instruction. J. dent. Educ., 1963, 27, 167-170.

Carpenter, C. R., & Greenhill, L. P. An investigation of closed-circuit television for teaching university courses. Instructional television research report no. 2, University Park, Pennsylvania: The Instructional Film Research Program, Pennsylvania State University, 1958.

Castle, C. H. Open-circuit television in postgraduate medical education. J. med. Educ., 1963, 38, 254-260.

Cochran, W. G., & Cox, Gertrude. Experimental designs. (2nd ed.) New York: John Wiley, 1957.

Comroe, J. H., Jr. The effects of direct chemical and electrical stimulation of the respiratory center in the cat. Amer. J. Physiol., 1943, 139, 490-497.

Cunningham, H. A. Lecture-demonstration v. individual laboratory method in science-teaching--a summary. Sci. Educ., 1946, 30, 70-79.

Dreher, R. E., & Beatty, W. H. Instructional television research. Project #1. San Francisco: San Francisco State College, 1958.

Dressel, P. L., & Mayhew, L. B. General education: exploration in evaluation. Washington: American Council on Education, 1954.

Eckert, Ruth E. Colleges and universities--programs. Encyclopedia of educational research. (3rd ed.) New York: MacMillan, 1960, Pp. 268-282.

Erickson, C. G., & Chausow, H. M. Chicago's tv college. Chicago: Chicago Public Schools, 1960.

Fisher, J. F., & Gershon-Cohen, J. Color translation and contrast enhancement of roentgenograms by television techniques. J. Einstein Med. Center, 1957, 5, 166-176.

Grant, T. S. Experimentation and development of new and more effective techniques and methods for training teachers in and presenting academic subject matter through the special media of television and video tapes in the health sciences. Revised application to the Commissioner of Education, U. S. Office of Education, Department of Health, Education, and Welfare, for funds to support a cooperative

research project under the provision of Public Law 864, 85th Congress (Title VII of the National Defense Education Act of 1958). Unpublished research application, Office of Television Research, University of California, San Francisco Medical Center, 1959.

Grant, T. S., Blancheri, R. L., Lorencki, S. F., & Merrill, I. R. Television in health sciences education: I. effectiveness of television within the dental laboratory. J. dent. Educ., 1962, 26, 146-151.

Green, E. J. The learning process and programmed instruction. New York: Holt, Rinehart and Winston, 1962.

Gropper, G. L., & Lumsdaine, A. A. An experimental comparison of a conventional tv lesson with a programmed tv lesson requiring active student response. Report #2. Studies in televised instruction. Pittsburgh: American Institute for Research, March, 1961.

Hard, W. Inter-connecting basic science and clinical facilities by television. Address at Fourth Annual Meeting, Council on Medical Television. Newsletter, June 6, 1962.

Hartman, F. R. Recognition learning under multiple channel presentation and testing conditions. AV Communication Rev., 1961, 9, 24-43.

Heyman, C. Chemoreceptors and regulation of respiration. Acta. Physiol. Scand., 1951, 22, 6-13.

Hovland, C. I., Lumsdaine, A. A., & Sheffield, F. P. Experiments in mass communication. Princeton: Princeton University, 1949.

Hoyt, C. Test reliability estimated by analysis of variance. Psychometrika, 1941, 6, 153-160.

Hurst, P. M., Jr. Relative effectiveness of verbal introductions to kinescope recordings and training films. Technical Report SDC 269-7-47. Port Washington, N. Y.: Special Devices Center, 1955.

Hutt, B. K., Horvath, S. M., & Spurr, C. B. Influence of varying degrees of passive limb movements in respiration and oxygen consumption of man. J. appl. Physiol., 1958, 12, 297-300.

Instructors manual for the test of science reasoning and understanding. Washington: American Council on Education (mimeographed), 1953.

Jackson, R. Learning from kinescopes and films. Technical Report SDC-20-TV-1. Port Washington, N. Y.: Special Devices Center, 1952.

Kanner, J. H. Teaching by television in the army--an overview. AV Communication Rev., 1958, 6, 172-188.

Kanner, J. H., Mindak, W. A., & Katz, S. A report on television in army training: the application of television and kinescope recordings to reduce instructor and student training time and training costs. Fort Gordon, Ga.: Office of the Chief Signal Officer, Department of the Army, June 1, 1958 (duplicated).

Kearney, N. C., & Cook, W. W. Curriculum. In C. W. Harris (Ed.) Encyclopedia of educational research, (3rd ed.) New York: MacMillan, 1960, Pp. 358-365.

Keele, C. A., & Neil, E. Sampson Wright's applied physiology (10th ed.). London: Oxford University Press, 1961.

Kendall, M. G. A course in multivariate analysis. New York: Hafner, 1961.

Lawshe, C. H., Jr. Principles of personnel testing. New York: McGraw Hill, 1948.

Lefkowith, E. F. The effect of pictorial stimuli similarity in teaching and testing. Unpublished doctoral dissertation, Pennsylvania State University, 1955.

Lepore, A. R., & Wilson, J. D. Instructional television research. Project #2. San Francisco: San Francisco State College, 1958.

Lumsdaine, A. A., May, M. A., & Hadsell, R. S. Questions spliced into a film for motivation and pupil participation. In A. A. Lumsdaine & M. A. May (Eds.) Learning from films. New Haven: Yale University, 1958, Pp. 72-83.

Macomber, F. G., & Siegel, L. Experimental study in instructional procedures. (Final Report). Oxford, Ohio: Miami University, 1960.

McGeoch, J. A., & Irion, A. L. The psychology of human learning. (2nd ed.) New York: Longmans, Green, 1952.

McGuire, F. L., Moore, F. J., & Harrison, C. A. The efficiency of television as applied to use of laboratory demonstrations in teaching. J. med. Educ., 1961, 36, 715-716.

McKeachie, W. J. Procedures and techniques of teaching: a survey of experimental studies. In N. Sanford (Ed.), The American college. New York: Wiley, 1962. Pp. 312-364.

Merrill, I. R. Television at the University of California Medical Center, annual report, 1961-1962. San Francisco: Office of Television Research, San Francisco Medical Center, University of California, June 30, 1962 (duplicated).

Michael, D. N. Some factors influencing the effects of audience participation on learning from a factual film. HRRL Memo Report #13A, Revised Report. Washington: U.S. Air Force, Human Resources Research Laboratories, December, 1951.

Miller, N. E. Graphic communication and the crisis in education. (Special issue AV Communication Rev., 1957, 5, 3) Washington, D.C.: Department of Audio-visual Instruction, National Education Association, 1957.

Olson, J. V. Dental education program at the University of Texas. J. Amer. dent. Ass., 1958, 56, 893-898.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. M. The measurement of meaning. Urbana: University of Illinois, 1957.

Owens, E. Intelligibility of words varying in familiarity. J. Speech & Hearing Res., 1961, 4, 113-129.

Paul, J., & Ogilvie, J. C. Mass media and retention. Explorations, 1955, 4, 120-123.

Postman, L., & Rosenzweig, M. R. Practice and transfer in the visual and auditory recognition of verbal stimuli. Amer. J. Psychol., 1956, 69, 209-226.

Postman, L., & Rosenzweig, M. R. Perceptual recognition of words. J. Speech & Hearing Disorders, 1957, 22, 245-253.

Rosenstein, A. J., & Kanner, J. H. Television and army training: color v. black and white. AV Communication Rev., 1961, 9, 44-49.

Roshal, S. M. Effects of learner representation in film-mediated perceptual motor learning. Technical Report SDC 269-7-5. Port Washington, N.Y.: Special Devices Center, 1949.

Ruhe, D. S., & Proud, G. O. Television: potentialities and problems in its use as a teaching medium. Trans. Amer. Acad. Ophthalmol., 1958, 62, 475-483.

Scheffe, H. The analysis of variance. New York: John Wiley, 1959.

Schramm, W. What we know about learning from instructional television. In Educational television, the next ten years. Stanford, California: The Institute for Communication Research, Stanford University, 1962, Pp. 52-76.

A staff summary and critique of a study on the Utah open-circuit teleclinics for practitioners. New York: Council on Medical Television of the Institute for Advancement of Medical Communication, March, 1961.

Tannenbaum, P. H. Instruction through television: a comparative study. Urbana: Institute of Communication Research, University of Illinois, 1956 (duplicated). (a)

Tannenbaum, P. H. Instruction through television: an experimental study. Urbana: Institute of Communication Research, University of Illinois, 1956 (duplicated). (b)

A test of science reasoning and understanding--biological sciences. Washington: American Council on Education, 1953.

Thorndike, E. L., & Lorge, I. The teacher's wordbook of 30,000 words. (2nd ptg.) New York: Teachers College, Columbia University, 1952.

Walker, Helen M., & Lev, J. Statistical inference. New York: Henry Holt, 1953.

Zworykin, V. K. Television techniques in biology and medicine. Advanc. in biol. & med. Physics, 1957, 5, 243-283.